Introduction

These Lecture Notes record the contents of a course I have taught occasionally over many years. The intent of the course is to provide graduate students, in Numerical Analysis or Scientific Computing, with the basic material from Functional Analysis needed for their Ph.D. work. For this reason, the material is always motivated by typical Numerical Analysis concerns. To give but three examples, topology is described in terms of neighborhood systems and convergence, completeness of a metric space is introduced in connection with fixed-point iteration, and the Hahn-Banach theorem is a natural outcome of trying to provide sharp bounds on the value of a linear functional at some function, given the value of certain linear functionals (e.g., point evaluations) at that function, along with some information about the size of that function.

In reading the material, carry along the various simple examples offered and test every concept and result on these, for a better understanding. Also, there are many problems embedded in the text and, ideally, the reader should try all of them as they occur. Solutions to almost all are available.

The first two chapters record basic facts from Linear Algebra and (Advanced) Calculus. For variety (and, possibly, a better understanding), some of these familiar facts are described in an unfamiliar way.

The next two chapters bring the basic facts about normed linear spaces and the representation of their (continuous) duals, including the Hahn-Banach Theorem and the corresponding practical problem of providing sharp bounds for the value of a linear functional on an element in terms of linear and nonlinear numerical information about that element (the Golomb-Weinberger interval).

Baire category is discussed in Chapter V, in connection with pointwise convergence, approximation order, uniform boundedness, and the open mapping/closed graph theorems.

This sets the stage for the remaining chapters, each written with some application(s) in mind.

Chapter VI brings the basic facts concerning convexity and uses these for the minimization of a convex functional over a convex set, including the characterization of best approximations from convex sets and, especially, linear subspaces in a nls.

Chapter VII discusses inner product spaces as nls's with a natural supply of continuous linear functionals. In this setting, earlier abstract results can be made concrete, leading to optimal interpolation (Synge's hypercircle method), including splines, reproducing kernels, and, ultimately, the Rayleigh-Ritz method for solving functional equations.

The basic idea of the Rayleigh-Ritz method is that of projecting an equation to be solved onto a suitable finite-dimensional subspace. The study of this method and its many variants leads to a discussion of compact perturbations of the identity, the topic of Chapter VIII.

The powerboundedness of a linear map is of basic concern in many contexts. In Chapter IX, it provides the motivation for a discussion of the spectrum of a linear map and that of 'nearby' maps.

Chapter X deals with the local approximation of maps by linear maps, with Newton's method and the Implicit Function theorem in Banach spaces as particular fruits of that discussion.

notation

For ready reference, I recall here the basic terms and notations concerning sets and maps as used in these lectures.

Sets and maps are the basic objects in Mathematics. The notations

$$\{x : P(x)\} \quad \text{and} \quad \{x \in X : P(x)\}$$

both describe the set of all elements x that have a certain property $P(x)$. In the first notation, the set of all x of interest here is to be understood from the context. Notations like $x \in X$, $Y \subset X$ or $Y \subseteq X$, $X \cap Y$, $X \cup Y$, are standard to indicate, respectively, that x is an element of X, Y is a subset of X, the intersection and the union of the two sets X and Y. I will use the notation $X\backslash Y$ (read: 'X **take away** Y', or, 'the **complement of** Y **in** X') for the set

$$X\backslash Y := \{x \in X : x \notin Y\}$$

of all elements of X that are not elements of Y, thus leaving me free to use the notation

$$X - Y := \{x - y : x \in X, y \in Y\}$$

for the difference of two subsets of a set in which subtraction is defined. In this connection, the asymmetric symbols ':=' and '=:' indicate that the expression on the colon's side is being defined by the expression on its other side. I will often write $X\backslash x$ rather than $X\backslash\{x\}$ for the set obtained by removing x from X (if possible). The cardinality of the set X is denoted by $\#X$.

If I need a notation for the collection of all subsets of the set S, I will use 2^S.

Specific sets used are: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ the set of, respectively, natural, integer, rational, real, and complex numbers. The base of the natural logarithm, $\mathrm{e} = 2.71828182\cdots$, and the imaginary unit, $\mathrm{i} := \sqrt{-1}$, deserve such special notation.

I will be careful to distinguish between sets and sequences, reserving braces, {, }, for the former, while using parentheses, (,), for the latter. In particular, {} denotes the empty set, while () denotes the empty sequence, and

$$x = (x(1), \ldots, x(n)) = (x(i) : i = 1, \ldots, n)$$

denotes an n-sequence, in X^n if each $x(i)$ is in X, or, more generally, in the **cartesian product**

$$X_1 \times \cdots \times X_n$$

of the sets $X_1, \ldots, X_n$ if $x(i) \in X_i$, all i.

The elements of $X^{m\times n}$ are, by definition, the (m, n)-**matrices** with entries from X. For $A \in X^{m\times n}$, I denote by $A(i, :)$ its ith **row**, i.e., the n-sequence $(A(i, 1), \ldots, A(i, n))$. Correspondingly, I denote by $A(:, j)$ its jth **column**, i.e., the m-sequence $(A(1, j), \ldots, A(m, j))$. Further, if $(a_1, \ldots, a_n)$ is a sequence of (real or complex) m-vectors, then I will use the nonstandard notation

$$[a_1, \ldots, a_n]$$

for the (real or complex) (m,n)-matrix whose jth column is a_j, all j. Its **transpose**

$$[a_1,\ldots,a_n]^t$$

is the (n,m)-matrix whose ith row is a_i, all i.

I will use $(a\,..\,b)$ to denote the open interval with endpoints a and b, reserving the notation (a,b) for the point in the plane with coordinates a and b or, more generally, the 2-sequence with entries a and b. In the same spirit, I will use $[a\,..\,b]$ to denote the closed interval with endpoint a and b, reserving the notation $[a,b]$ for the matrix with columns a and b.

The **function** or **map**

$$f : T \to U : t \mapsto f(t)$$

carries the typical element t of its **domain** $T =: \operatorname{dom} f$ to its corresponding **value** $f(t)$ in its **target** $U =: \operatorname{tar} f$. Its **range** (or, **image**) is

$$\operatorname{ran} f := \{f(t) : t \in T\} =: f(T),$$

while, if $\operatorname{tar} f$ is some field or, more generally, some linear space, then

$$\operatorname{supp} f := \{x \in \operatorname{dom} f : f(x) \neq 0\}$$

denotes its **support**. More generally, $f(S) := \{f(s) : s \in S\}$ is the **image** of $S \subset T$ under f, while $f^{-1}(W) := \{t \in T : f(t) \in W\}$ is the **preimage** of $W \subset U$ under f. Such f is **1-1** (or, **one-one**, **injective**, or a **monomorphism**) in case $f(t) = f(t') \implies t = t'$. Such f is **onto** (or **surjective**, or an **epimorphism**) in case $f(T) = U$. These notions correspond to uniqueness, resp. existence of a solution to the equation $f(?) = u$ for every $u \in U$. Such f is **invertible** in case it is both 1-1 and onto. In that case, there exists $g : U \to T$, necessarily unique, so that $fg = 1$ as well as $gf = 1$. We denote such g by f^{-1} and call it the **inverse** of f. Here, 1 denotes the identity map, on whatever set the context indicates. More generally, if $fg = 1$, then g is a **right inverse** for f and f is a **left inverse** for g and, in this situation, f is onto and g is 1-1. Conversely, if f is $\begin{smallmatrix}\text{1-1}\\ \text{onto}\end{smallmatrix}$, then it has a $\begin{smallmatrix}\text{left}\\ \text{right}\end{smallmatrix}$ inverse.

Map **composition** is used here. If $f : T \to U$ and $g : U \to W$, then $gf : T \to W : t \mapsto g(f(t))$. The inversion of order here is unavoidable since we write the function symbol on the left, yet usually write maps as mapping from left to right. This is particularly awkward since map composition, although **associative** (i.e., $h(gf) = (hg)f$), fails to be commutative.

Strictly speaking, two maps are **equal** iff they have the same domain and target and agree on their common domain. Thus, if $f : T \to U$ has its range in some proper subset W of U, then the map

$$f|^W : T \to W : t \mapsto f(t)$$

is different from f. Finally, if $S \subset T$, then the **restriction** of $f : T \to U$ to S is

$$f|_S : S \to U : s \mapsto f(s).$$

The set of all maps from T to U is denoted by

$$U^T := \{f : T \to U\}.$$

In this sense, 2^T is short-hand for $\{0,1\}^T$ as one identifies the subsets S of T with their **characteristic function**

$$\chi_S : T \to \{0,1\} : t \mapsto \begin{cases} 1 & t \in S; \\ 0 & \text{otherwise.} \end{cases}$$

Also

$$U^n := U^{\{1,\ldots,n\}} := \{(u(1),\ldots,u(n)) : u(j) \in U\}.$$

I think it important to distinguish between the map f and its value $f(t)$ at some point t. This makes it necessary to invent a name for the **power function**. I will use the symbol $()^r$ for it (but will *not* usually write $()^r(t)$ for its value at t :-). The definition in full:

$$()^r : \mathbb{R} \to \mathbb{R} : t \mapsto t^r.$$

Occasionally, there will be use for its d-variate version:

$$()^\alpha := \prod_{j=1}^{d} ()^{\alpha(j)} : x \in \mathbb{R}^d \mapsto x(1)^{\alpha(1)} \cdots x(d)^{\alpha(d)}, \quad \alpha \in \mathbb{Z}_+^d := \{\alpha \in \mathbb{Z}^d : \alpha(j) \ge 0, \text{all } j\}.$$

I will also use the convenient **place holder** notation to indicate the map $f(\cdot,u)$ on some set T obtained from some map $f : T \times U \to W$ by fixing a point $u \in U$. Explicitly,

$$f(\cdot,u) : T \to W : t \mapsto f(t,u).$$

For example, $(\cdot - a)(\cdot - b)\cdots(\cdot - z)$ is the polynomial with leading coefficient 1 that vanishes at the points $a, b, \ldots, z$.

Another non-standard notation was used earlier when I talked about the equation $f(?) = u$, thus using the question mark to indicate that the equation is to be solved and exactly where the unknown occurs in the equation.

While the notation $\sup S := \sup\{s : s \in S\}$ for the supremum or least upper bound of a real set S is standard (along with its counterpart, $\inf S$), the notation

$$\operatorname{argmax} f(M) := \{m \in M : f(m) = \sup f(M)\}$$

for M a subset of the domain of the real-valued map f may be less familiar (though no less useful). I will even use the incorrect but suggestive notation

$$m = \operatorname{argmax} f(M)$$

to mean that $\operatorname{argmax} f(M) = \{m\}$, and use argmin analogously.

The quantifiers ($\exists$, $\forall$) are usually followed by a braces-enclosed list of one or more expressions. For example,

$$\forall\{f\}\, \exists\{g\} \quad f + g = 0$$

is to be read '*for all f, there exists g such that $f + g = 0$*'.

The text is 'font-coded' in the following way. Words emphasized appear in *italics*, terms being defined appear in **boldface**, and all formal statements are set in a *slanted font*.

