NOTES ON NUMERICAL ANALYSIS

YANNAN MAO

1ST JUNE 2024

CONTENTS

APPROXIMATION THEORY

1.1 REVIEW OF RELEVANT TOPICS

An ***algebraic structure*** consists of one or more sets with one or more binary operations thereon. We denote an algebraic structure by an n-tuple whose first element is the ***carrier set of the algebraic structure*** and is followed by other sets followed by the binary operations.

When there is no confusion, we refer to the algebraic structure by the carrier set thereof.

1.1.1 GROUPS, RINGS, AND FIELDS

DEFINITION 1. A ***group*** is an ordered pair $(G, *)$ wherein

(a) G is a set; and

(b) $*$ is a binary operation on G which satisfies

 (i) $*$ is closed;

 (ii) $*$ is associative;

 (iii) there exists an identity element for $*$; and

 (iv) for each each element in G there exists an inverse element thereof.

If $*$ is commutative, the group is ***abelian***.

DEFINITION 2. A ***ring*** is an ordered triple $(K, *, \star)$ wherein

(a) $(K, *)$ is an abelian group; and

(b) $\star$ is a binary operation on K which satisfies

 (i) $\star$ is closed;

 (ii) $\star$ is commutative;

 (iii) $\star$ is associative; and

 (iv) $\star$ is distributive over $*$.

If there exists an identity element for $\star$, the ring is ***unital***.

If $\star$ is commutative, the ring is ***commutative***.

DEFINITION 3. A ***field*** is a unital commutative ring $(F, *, \star)$ which satisfies

(a) the identity for $*$ is distinct from the identity for $\star$;
(b) for each element in F distinct from the identity for $*$ there exists an inverse element thereof for $\star$.

The set $\mathbb{Z}$ with addition and multiplication is a unital commutative ring which is not a field, referred to as the ***ring of integers***.

The set $\mathbb{C}$ with addition and multiplication is a field. A ***subfield of the field*** $\mathbb{C}$ is a set $F \subseteq \mathbb{C}$ which is itself a field with addition and multiplication.

The set $\mathbb{Q}$ is a subfield of $\mathbb{C}$. Any subfield of $\mathbb{C}$ contains $\mathbb{Q}$.

If $(K, *, \star)$ is a ring, we may apply $*$ to the identity for $\star$ with itself finitely many times and obtain the identity for $*$. If this happens in K, the least positive integer n such that applying $*$ to the identity for $\star$ n times results in the identity for $*$ is the ***characteristic of the ring*** K. If this does not happen in K, then K is a ***ring of characteristic zero***.

Any subfield of $\mathbb{C}$ is of characteristic zero.

1.1.2 VECTOR SPACES

Suppose $(F, +, \cdot)$ is a field. We write $\lambda\mu$ as a shorthand for $\lambda \cdot \mu$ for any $\lambda, \mu \in F$. We denote by 0 the identity for $+$ in F, and 1 the identity for $\cdot$ in F. We refer to 0 as ***zero***.

DEFINITION 4. A ***vector space*** is an ordered quadruple $(V, F, +, \cdot)$ wherein

(a) V is a set, whose elements are referred to as ***vectors***;
(b) F is a field, whose elements are referred to as ***scalars***;
(c) $+$ is a binary operation on V, referred to as ***vector addition***, which satisfies
    (i) $+$ is closed;
    (ii) $+$ is commutative;
    (iii) $+$ is associative;
    (iv) there exists an identity element for $+$, referred to as the ***zero vector*** and denoted by $\mathbf{0}$; and

(v) for each $\boldsymbol{a} \in V$ there exists an inverse element thereof, denoted by $-\boldsymbol{a}$;

and

(d) $\cdot : F \times V \to V$ is a binary operation, referred to as ***scalar multiplication***, which satisfies

(i) $1 \cdot \boldsymbol{a}$ for each $\boldsymbol{a} \in V$;

(ii) $(\lambda\mu) \cdot \boldsymbol{a} = \lambda \cdot (\mu \cdot \boldsymbol{a})$ for any $\lambda, \mu \in F$ and $\boldsymbol{a} \in V$;

(iii) $\lambda \cdot (\boldsymbol{a} + \boldsymbol{b}) = \lambda \cdot \boldsymbol{a} + \lambda \cdot \boldsymbol{b}$ for any $\lambda \in F$ and $\boldsymbol{a}, \boldsymbol{b} \in V$; and

(iv) $(\lambda + \mu) \cdot \boldsymbol{a} = \lambda \cdot \boldsymbol{a} + \mu \cdot \boldsymbol{a}$ for any $\lambda, \mu \in F$ and $\boldsymbol{a} \in V$.

For vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we refer to $\boldsymbol{a} + \boldsymbol{b}$ as the ***sum of a and b***. For a scalar λ and a vector $\boldsymbol{a}$, we refer to $\lambda \cdot \boldsymbol{a}$ as the ***product of*** λ ***and a***.

We write $\lambda\boldsymbol{a}$ as a shorthand for the product of λ and $\boldsymbol{a}$.

The same set of vectors V may be the carrier set of a number of distinct vector spaces. When it is desirable to specify the field, we say V is a ***vector space over the field*** F.

A ***subspace of a vector space*** V is a set $W \subseteq V$ which is itself a vector space over the same field with the operations of vector addition and scalar multiplication on V.

1.1.3 BASES AND DIMENSIONS

Suppose F is a field.

DEFINITION 5. Let V be a vector space over F. A set $S \subseteq V$ is ***linearly dependent*** if there exist $(n+1)$ distinct vectors $\boldsymbol{a}_0, \dots, \boldsymbol{a}_n \in S$ and $(n+1)$ scalars $\lambda_0, \dots, \lambda_n \in F$, not all of which are zero, such that

$$\sum_{i=0}^{n} \lambda_i \boldsymbol{a}_i = \boldsymbol{0}.$$

A set which is not linearly dependent is ***linearly independent***.

If the set S is finite, we also say that the vectors in S are linearly dependent or independent.

A set which contains a linearly dependent set is linearly dependent.

A subset of a linearly independent set is linearly independent.

A set which contains **0**, the zero vector, is linearly dependent.

A set is linearly independent if and only if each finite subset thereof is linearly independent.

DEFINITION 6. Let V be a vector space, and let $S \subseteq V$. The ***subspace spanned by the set*** S is the intersection W of all subspaces of V which contain S.

If S is finite, we say the vectors in S span W.

DEFINITION 7. Let V be a vector space. A ***basis for*** V is a linearly independent set of vectors in V which spans V.

A vector space is ***finite-dimensional*** if it has a finite basis, and ***infinite-dimensional*** otherwise.

It follows from the axiom of choice that every vector space has a basis.

If V is a finite-dimensional vector space, any two bases for V are of the same cardinality. This allows us to define the ***dimension of a finite vector space*** as the cardinality of any basis therefor.

1.1.4 THE n-TUPLE SPACE AND THE STANDARD BASIS THEREOF

Suppose F is a field.

Example 1. We denote by F^n the set of all n-tuples $\boldsymbol{a} = (a_0, \dots, a_{n-1})$ of scalars $a_i \in F$, $0 \leq i < n$.

If $\boldsymbol{a}, \boldsymbol{b} \in F$ with $\boldsymbol{a} = (a_0, \dots, a_{n-1})$ and $\boldsymbol{b} = (b_0, \dots, b_{n-1})$, we define the sum of $\boldsymbol{a}$ and $\boldsymbol{b}$ by

$$\boldsymbol{a} + \boldsymbol{b} = (a_0 + b_0, \dots, a_{n-1} + b_{n-1}).$$

And we define the product of $\lambda \in F$ and $\boldsymbol{a}$ by

$$\lambda \boldsymbol{a} = (\lambda a_0, \dots, \lambda a_{n-1}).$$

Thus defined, F^n is a vector space, referred to as the *n-**tuple space***.

As a special case for when $n = 1$, the field F itself is a vector space.

Example 2. Let $E \subseteq F^n$ consist of the vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \dots, \boldsymbol{e}_{n-1}$ defined by

$$\begin{aligned}
\boldsymbol{e}_0 &= (1, 0, \dots, 0), \\
\boldsymbol{e}_1 &= (0, 1, \dots, 0), \\
&\vdots \\
\boldsymbol{e}_{n-1} &= (0, 0, \dots, 1).
\end{aligned}$$

Thus defined, the set E is the ***standard basis of the** n-**tuple space*** F^n.

1.1.5 ALGEBRAS

Suppose F is a field.

DEFINITION 8. An ***algebra over the field*** F is an ordered quintuple $(\mathcal{A}, F, +, \cdot, \times)$ wherein

(a) $(\mathcal{A}, F, +, \cdot)$ is a vector space; and

(b) $\times$ is a binary operation on $\mathcal{A}$, referred to as ***vector multiplication***, which satisfies

  (i) $\times$ is closed;

  (ii) $\times$ is distributive over $+$; and

  (iii) $\lambda(\boldsymbol{a} \times \boldsymbol{b}) = (\lambda \boldsymbol{a}) \times \boldsymbol{b} = \boldsymbol{a} \times (\lambda \boldsymbol{b})$ for any $\lambda \in F$ and $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{A}$.

If there exists an identity element for $\times$, the algebra is ***unital***.

If $\times$ is commutative, the algebra is ***commutative***.

If $\times$ is associative, the algebra is ***associative***.

For vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we refer to $\boldsymbol{a} \times \boldsymbol{b}$ as the ***product of a and b***.

Example 3. The set of $(n \times n)$ matrices over the field F, with the usual matrix operations, is a unital associative algebra, which is not commutative if $n > 1$.

As a special case for when $n = 1$, the field F itself is a unital commutative associative algebra.

1.1.6 POLYNOMIALS

Suppose F is a field.

Example 4. Let S be a nonempty set, and let V be the set of functions from S into the field F.

We define the sum of $f, g \in V$ by

$$(f + g)(s) = f(s) + g(s)$$

for each $s \in S$. And we define the product of $\lambda \in F$ and $f \in V$ by

$$(\lambda f)(s) = \lambda f(s)$$

for each $s \in S$.

Thus defined, V is a vector space, referred to as the ***space of functions from the set*** S ***into the field*** F.

Note that F^n, the n-tuple space, is a special case of the space of functions, if we take the set S to be $\mathbb{N}_{<n}$, the set of natural numbers less than n, and consider an n-tuple to be a function indexed by $\mathbb{N}_{<n}$.

We now consider the space of functions from $\mathbb{N}$ into F. We denote this space by F^∞. A vector in F^∞ is thus an infinite sequence $f = (f_0, f_1, f_2, \ldots)$ of scalars $f_i \in F$, $i \in \mathbb{N}$.

For two vectors $f, g \in F^\infty$ with $f = (f_0, f_1, f_2, \ldots)$ and $g = (g_0, g_1, g_2, \ldots)$, we define $fg = f \times g$, the product of f and g, by

$$(fg)_n = f_0 g_n + f_1 g_{n-1} + f_2 g_{n-2} + \cdots + f_n g_0 = \sum_{i=0}^{n} f_i g_{n-i}.$$

We may verify that F^∞ is an algebra with vector multiplication as defined above.

Note that, thus defined, vector multiplication is commutative and associative, and the vector $1 = (1, 0, 0, \ldots)$ is the vector-multiplicative identity. Therefore, F^∞ is a unital commutative associative algebra over the field F.

The vector $(0, 1, 0, \ldots)$ plays a distinguished role in F^∞, and we denote it by x. The product of x multiplied by itself n times is denoted by x^n, and we put $x^0 = 1$. Then, the vector x^n is given by

$$x_i^n = \begin{cases} 1, & \text{if } i = n, \\ 0, & \text{otherwise,} \end{cases}$$

wherein x_i^n denotes the i-th entry of x^n.

We observe that the set $\{x^n : n \in \mathbb{N}\} \subseteq F^\infty$ is both independent and infinite. Thus, the space F^∞ is infinite-dimensional.

The algebra F^∞ is referred to as the ***algebra of formal power series over*** F. A vector $f = (f_0, f_1, f_2, \ldots) \in F^\infty$ is written

$$f = f_0 + f_1 x + f_2 x^2 + \cdots = \sum_{n=0}^{\infty} f_n x^n.$$

DEFINITION 9. Let $F[x]$ be the subspace of F^∞ spanned by $\{x^n : n \in \mathbb{N}\}$.

Thus defined, a vector in $F[x]$ is referred to as a ***polynomial over the field*** F.

A nonzero vector $f \in F^\infty$ is a polynomial if and only if there exists an $n \in \mathbb{N}$ such that $f_n \neq 0$ and that $f_i = 0$ for each $i > n$. This natural number n is unique, referred to as the ***degree of the polynomial*** f and denoted by $\deg f$.

The degree of the zero vector $(0, 0, 0, \ldots)$, referred to as the ***zero polynomial*** and denoted by 0, is undefined.

If $f \in F[x]$ is a nonzero polynomial of degree n, it follows that

$$f = \sum_{i=0}^{n} f_i x^i, \quad f_n \neq 0.$$

Let f and g be nonzero polynomials over F. Then fg is also a nonzero polynomial and

$\deg(fg) = \deg f + \deg g$. If in addition $f + g \neq 0$, then $\deg(f + g) \leq \max\{\deg f, \deg g\}$.

1.1.7 ROOTS OF POLYNOMIALS

Suppose F is a field.

Lemma 1. *Let f and d be nonzero polynomials over the field F such that $\deg f \geq \deg d$. Then there exists a polynomial $g \in F[x]$ such that either $f - dg = 0$ or $\deg(f - dg) < \deg f$.*

Proof. Suppose that

$$f = f_m x^m + \sum_{i=0}^{m-1} f_i x^i, \quad f_m \neq 0,$$

and that

$$d = d_n x^n + \sum_{j=0}^{n-1} d_j x^j, \quad d_n \neq 0.$$

Then $m \geq n$, and either

$$f - \frac{f_m}{d_n} x^{m-n} = 0$$

or

$$\deg\left(f - \frac{f_m}{d_n} x^{m-n} d\right) < \deg f.$$

Hence, we may take $g = \dfrac{f_m}{d_n} x^{m-n}$, which satisfies the condition. □

Theorem 1. *If f and d are polynomials over the field F and $d \neq 0$, then there exist a unique pair of polynomials $q, r \in F[x]$ such that*

(a) $f = dq + r$; and

(b) either $r = 0$ or $\deg r < \deg d$.

Proof. If $f = 0$ or $\deg f < \deg d$, then $q = 0$ and $r = f$.

If $f \neq 0$ and $\deg f \geq \deg d$, by lemma 1 we may choose a polynomial g_0 such that $f - dg_0 = 0$ or $\deg(f - dg_0) < \deg f$. If $f - dg_0 = 0$, then $q = g_0$ and $r = 0$. If $f - dg_0 \neq 0$ and $\deg(f - dg_0) < \deg d$, then $q = g_0$ and $r = f - dg_0$.

If $f - dg_0 \neq 0$ and $\deg(f - dg_0) \geq \deg d$, we apply lemma 1 again and choose another polynomial g_1 such that $f - d(g_0 + g_1) = 0$ or $\deg\big(f - d(g_0 + g_1)\big) < \deg(f - dg_0)$. Continuing this process, in the end we obtain polynomials q and r such that $r = 0$ or $\deg r < \deg d$ and that $f = dq + r$.

To prove that q and r are unique, suppose for the sake of contradiction we also have $f = dq' + r'$ with $r' = 0$ or $\deg r' < \deg d$. It follows that $dq + r = dq' + r'$, and then $d(q - q') = r' - r$. Because $q - q' \neq 0$, we have $d(q - q') \neq 0$, and so $\deg d + \deg(q - q') = \deg(r' - r)$. This is a contradiction, as $\deg(r' - r) < \deg d$. Ergo, our supposition is false.

Hence, the theorem holds. □

DEFINITION 10. Let d be a nonzero polynomial over the field F. If $f \in F[x]$, then by theorem 1 there exists at most one polynomial $q \in F[x]$ such that $f = dq$. If such a q exists, we say that d ***divides*** f, that f ***is divisible by*** d, that d is a ***divisor of*** f, that f is a ***multiple of*** d, and that q is the ***quotient of*** f ***and*** d; and we write $d \mid f$ and $q = f/d$.

DEFINITION 11. Let f be a polynomial over the field F. A scalar $c \in F$ is a ***root of*** f ***over the field*** F if $f(c) = 0$.

Corollary 1. *Let f be a polynomial over the field F, and let $c \in F$. Then $(x - c) \mid f$ if and only if c is a root of f.*

Proof. By theorem 1, $f = (x - c)q + r$ wherein r is a scalar polynomial. Then, $f(c) = r(c)$. It follows that $r = 0$ if and only if $f(c) = 0$. □

If c is a root of a polynomial f, the ***multiplicity of*** c ***as a root of*** f is the largest positive integer n such that $(x - c)^n \mid f$.

The multiplicity of any root of f is at most $\deg f$.

Thus concludes our review of relevant topics.

1.2 BÉZIER CURVES

We investigate polynomials in the d-dimensional Euclidean space $\mathbb{R}^d$.

DEFINITION 12. A ***polynomial of degree*** n ***in*** $\mathbb{R}^d$ is a function $\boldsymbol{p} : \mathbb{R} \to \mathbb{R}^d$ defined by

$$\boldsymbol{p}(t) = \sum_{i=0}^{n} \boldsymbol{a}_i t_i, \quad \boldsymbol{a}_i \in \mathbb{R}^d \text{ and } \boldsymbol{a}_n \neq \mathbf{0}.$$

We denote the space of polynomials of degree less than or equal to n in $\mathbb{R}^d$ by P_n^d.

As a special case for when $d = 1$, the space is denoted by P_n and a polynomial $p \in P_n$ is given by

$$p(t) = \sum_{i=0}^{n} a_i t_i, \quad a_i \in \mathbb{R}.$$

Let $\{\boldsymbol{e}_0, \dots, \boldsymbol{e}_{d-1}\}$ be the standard basis of $\mathbb{R}^d$. If $\{p_0, \dots, p_n\}$ is a basis for P_n, then the polynomials

$$\{p_i \boldsymbol{e}_j : i \in \mathbb{N}_{\leq n} \text{ and } j \in \mathbb{N}_{<d}\}$$

form a basis for P_n^d.

1.2.1 BERSTEIN BASIS POLYNOMIALS

We denote the closed interval between any $a, b \in \mathbb{R}$ by $[a, b]$. Namely,

$$[a, b] = \{x : x = \lambda a + (1 - \lambda) b \text{ and } 0 \leq \lambda \leq 1\}.$$

DEFINITION 13. The ***i-th Bernstein basis polynomial of degree*** n is the polynomial $B_{i,n} \in P_n$ defined by

$$B_{i,n}(t) = \binom{n}{i} (1 - t)^{n-i} t^i$$

for each $t \in [0, 1]$, wherein $i \in \mathbb{N}_{\leq n}$.

Example 5. The first few Bernstein polynomials are

$$
\begin{aligned}
&B_{0,0}(t) = 1, \\
&B_{0,1}(t) = 1 - t, \qquad B_{1,1}(t) = t, \\
&B_{0,2}(t) = (1-t)^2, \quad B_{1,2}(t) = 2t(1-t), \quad B_{2,2}(t) = t^2, \\
&B_{0,3}(t) = (1-t)^3, \quad B_{1,3}(t) = 3t(1-t)^2, \quad B_{2,3}(t) = 3t^2(1-t), \quad B_{3,3}(t) = t^3.
\end{aligned}
$$

For any closed interval $[a, b]$, we can apply to it an affine transformation onto the unit interval $[0, 1]$ by

$$t \mapsto \frac{t-a}{b-a}.$$

Then, we define Bernstein basis polynomials with respect to any closed interval $[a, b]$ by

$$B_{i,n}(t; a, b) = B_{i,n}\left(\frac{t-a}{b-a}\right) = \frac{1}{(b-a)^n}\binom{n}{i}(t-a)^i(b-t)^{n-i}$$

for each $t \in [a, b]$, wherein $i \in \mathbb{N}_{\le n}$.

Theorem 2. *The Bernstein basis polynomials $B_{i,n}$ satisfy the following properties:*

(a) 0 *is a root of multiplicity i of $B_{i,n}$.*

(b) 1 *is a root of multiplicity $(n-i)$ of $B_{i,n}$.*

(c) $B_{i,n}(t) = B_{n-i,n}(1-t)$ *for each* $t \in [0, 1]$.

(d) $(1-t)B_{0,n} = B_{n,n+1}$ *for each* $t \in [0, 1]$.

(e) $tB_{n,n} = B_{n+1,n+1}$ *for each* $t \in [0, 1]$.

(f) The $(n+1)$ Bernstein basis polynomials $B_{i,n}$, $i \in \mathbb{N}_{\le n}$, are nonnegative on $[0, 1]$ *and form a partition of unity; i.e.,* $B_{i,n}(t) \ge 0$ *and*

$$\sum_{i=0}^{n} B_{i,n}(t) = 1$$

for each $t \in [0, 1]$.

(g) $B_{i,n}$ *has a unique local maximum at* $\frac{i}{n}$.

(h) The recurrence relation

$$B_{i+1,n}(t) = tB_{i,n-1}(t) + (1-t)B_{i+1,n-1}(t)$$

for each $t \in [0,1]$ *and* $i \in \mathbb{N}_{<n}$ *holds.*

(i) The $(n+1)$ *Bernstein basis polynomials* $B_{i,n}$, $i \in \mathbb{N}_{\le n}$, *form a basis for* P_n.

Proof. TODO □

Because the $(n+1)$ Bernstein basis polynomials $B_{i,n}$, $i \in \mathbb{N}_{\le n}$, form a basis for P_n, the polynomials

$$\{B_{i,n}\boldsymbol{e}_j \,:\, i \in \mathbb{N}_{\le n} \text{ and } j \in \mathbb{N}_{<d}\}$$

form a basis for P_n^d. Thus, a polynomial $\boldsymbol{p} \in P_n^d$ may be written

$$\boldsymbol{p} = \sum_{i=0}^{n} B_{i,n}\boldsymbol{\beta}_i, \quad \boldsymbol{\beta}_i \in \mathbb{R}^d,$$

wherein each coefficient $\boldsymbol{\beta}_i$, $i \in \mathbb{N}_{\le n}$, is given by

$$\boldsymbol{\beta}_i = \sum_{j=0}^{d-1} \lambda_j \boldsymbol{e}_j$$

for some $\lambda_j \in \mathbb{R}$, $j \in \mathbb{N}_{<d}$. The coefficients $\boldsymbol{\beta}_j$ are referred to as ***Bézier coefficients*** and the ***control points of*** $\boldsymbol{p}$.