A3

An Interpolation Theorem

Overview

Appendix 3 presents an interpolation theorem for linear functions that is used in Section 7.1. From this Haar's characterization of Chebyshev spaces follows, as do alternate proofs of many of the basic inequalities.

An Interpolation Theorem

The main result of this section, Theorem A.3.3, is an interpolation theorem that plays an important role in Section 7.1. Further applications are given in the exercises.

Throughout this appendix we use the following notation. Let Q be a compact Hausdorff space. Let $C(Q)$ be the space of all real- or complex-valued continuous functions defined on Q. Let P be a (usually finite-dimensional) linear subspace of $C(Q)$ over $\mathbb{R}$ if $C(Q)$ is real or over $\mathbb{C}$ if $C(Q)$ is complex. The function $f \in C(Q)$ is said to be *orthogonal* to P, written as $f \perp P$, if

$$\|f\|_Q \leq \|f + p\|_Q \qquad \text{for all } p \in P\,.$$

This is an L_∞ analog of the more usual L_2 notion of orthogonality, as in Section 2.2. The following two lemmas give necessary and sufficient conditions for the relation $f \perp P$.

Lemma A.3.1. *Let $0 \not\equiv f \in C(Q)$. The function f is orthogonal to P if and only if there exists no $p \in \mathcal{P}$ such that*

$$\mathrm{Re}\big(\overline{f(x)}p(x)\big) > 0 \tag{A.3.1}$$

holds on

$$E := E(f) := \{x \in Q : |f(x)| = \|f\|_Q\}\,. \tag{A.3.2}$$

Proof. Assuming there exists $p \in \mathcal{P}$ satisfying (A.3.1) on E defined by (A.3.2), we show that

$$\|f - \epsilon p\|_Q < \|f\|_Q$$

for some $\epsilon > 0$. Since the set E defined by (A.3.2) is compact, $\mathrm{Re}\big(\overline{f(x)}p(x)\big)$ attains its positive minimum, say, $2\delta > 0$, on E, and there exists an open set G containing E such that

$$\mathrm{Re}\big(\overline{f(x)}p(x)\big) > \delta > 0\,, \qquad x \in G\,.$$

Since $G^c := Q \setminus G$ is also compact, there exists an $\alpha > 0$ such that

$$|f(x)| < (1-\alpha)\|f\|_Q\,, \qquad x \in G^c\,.$$

Thus, with a sufficiently small $\epsilon > 0$,

$$|f(x) - \epsilon p(x)|^2 \le \|f\|_Q^2 - 2\epsilon\delta + \epsilon^2\|p\|_Q^2 < \|f\|_Q^2\,, \qquad x \in G\,,$$

while

$$|f(x) - \epsilon p(x)| \le (1-\alpha)\|f\|_Q + \epsilon\|p\|_Q < \|f\|_Q\,, \qquad x \in G^c\,.$$

Therefore

$$\|f - \epsilon p\|_Q < \|f\|_Q$$

if $\epsilon > 0$ is small enough.

Conversely, if f is not orthogonal to P, then there exists $p \in P$ such that $\|f - p\|_Q^2 < \|f\|_Q^2$, so

$$2\,\mathrm{Re}\big(\overline{f(x)}p(x)\big) > \|p\|_Q^2 \ge 0\,, \qquad x \in E\,,$$

and the proof is finished. □

Lemma A.3.2. *Let $0 \neq f \in C(Q)$ and let P be an n-dimensional linear subspace of $C(Q)$ over $\mathbb{R}$ if $C(Q)$ is real or over $\mathbb{C}$ if $C(Q)$ is complex. Then the function f is orthogonal to P if and only if there exist points $x_1, x_2, \dots, x_r$ in $E(f)$ defined by (A.3.2) and positive real numbers $c_1, c_2, \dots, c_r$, where $1 \le r \le n+1$ when $C(Q)$ is real, and $1 \le r \le 2n+1$ when $C(Q)$ is complex, such that*

$$\sum_{i=1}^{r} c_i \overline{f(x_i)} p(x_i) = 0\,, \qquad p \in P\,. \tag{A.3.3}$$

Proof. Suppose (A.3.3) holds with some positive real $c_1, c_2, \dots, c_r$ satisfying $\sum_{i=1}^{r} c_i = 1$. As $|f(x_i)| = \|f\|_Q$, we have

$$\begin{aligned} \|f\|_Q^2 &= \sum_{i=1}^{r} c_i \overline{f(x_i)} f(x_i) = \sum_{i=1}^{r} c_i \overline{f(x_i)} (f(x_i) - p(x_i)) \\ &\le \|f\|_Q \sum_{i=1}^{r} c_i \max_{1 \le j \le r} |f(x_j) - p(x_j)| \le \|f\|_Q \|f - p\|_Q \end{aligned}$$

for every $p \in P$, so $f \perp P$. (Note that $r \le n+1$ or $r \le 2n+1$, respectively, was not needed for this part of the proof, so the sufficiency of (A.3.3) is valid with no hypothesis about r.)

Conversely, suppose $f \perp P$. Let $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ be a basis for P over $\mathbb{R}$ (or $\mathbb{C}$, respectively), and consider the map $T : Q \to \mathbb{R}^n$ (or $\mathbb{C}^n$, respectively) defined by

$$T(x) = \overline{f(x)} (\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x))\,.$$

Observe that the origin is in the convex hull of

$$T(E) := \{T(x) : x \in E\}\,,$$

otherwise by the principle of separating hyperplanes (a corollary of the Hahn-Banach theorem; see Rudin [73]), there would exist complex numbers $a_1, a_2, \dots, a_n$ such that

$$\operatorname{Re}\left(\sum_{i=1}^{n} a_i \overline{f(x)} \varphi_i(x) \right) > 0\,, \qquad x \in E\,.$$

Hence, with $p := \sum_{i=1}^{n} a_i \varphi_i \in P$,

$$\operatorname{Re}\big(\overline{f(x)} p(x)\big) > 0\,, \qquad x \in E$$

and f is not orthogonal to P by Lemma A.3.1, which contradicts our assumption.

Now it follows from Caratheodory's lemma (see E.1) that there exist points $x_1, x_2, \ldots, x_r$ in E and positive real numbers $c_1, c_2, \ldots, c_r$, where $1 \le r \le n+1$ when $C(Q)$ is real and $1 \le r \le 2n+1$ when $C(Q)$ is complex, such that

$$\sum_{i=1}^{r} c_i \overline{f(x_i)} \varphi_k(x_i) = 0, \qquad k = 1, 2, \ldots, n.$$

Therefore

$$\sum_{i=1}^{r} c_i \overline{f(x_i)} p(x_i) = 0, \qquad p \in P,$$

and the lemma is proved. □

Theorem A.3.3 (Interpolation of Linear Functionals). *Let $C(Q)$ be the set of real- (complex-) valued continuous functions on the compact Hausdorff space Q. Let P be an n-dimensional linear subspace of $C(Q)$ over $\mathbb{R}$ ($\mathbb{C}$). Let $L \neq 0$ be a real- (complex-) valued linear functional on P. Then there exist points $x_1, x_2, \ldots, x_r$ in Q, and nonzero real (complex) numbers $a_1, a_2, \ldots, a_r$, where $1 \le r \le n$ in the real case and $1 \le r \le 2n-1$ in the complex case, such that*

$$L(p) = \sum_{i=1}^{r} a_i p(x_i), \qquad p \in P \tag{A.3.4}$$

and

$$\|L\| = \sum_{i=1}^{r} |a_i|, \tag{A.3.5}$$

where

$$\|L\| := \sup\{|L(p)| : p \in P, \ \|p\|_Q \le 1\}.$$

Proof. Because of the finite dimensionality of P, there exists an element $p^* \in P$ (called an *extremal element* for L) such that $\|p^*\|_Q = 1$ and $L(P^*) = \|L\|$. Let P_0 denote the null-space of L, so

$$P_0 := \{p \in P : L(p) = 0\}.$$

Now p^* is orthogonal to P_0 because if

$$\|p^* + p_0\|_Q < \|p^*\|_Q = 1$$

for some $p_0 \in P$, then $g := p^* + p_0$ satisfies $\|g\|_Q < 1$ and $L(g) = \|L\|$, which is impossible. Note that the dimension of P_0 over $\mathbb{R}$ is $n-1$ in the real case and $2n-2$ in the complex case. So by Lemma A.3.2 there exist points $x_1, x_2, \ldots, x_r$ in

$$E := \{x \in Q : |p^*(x)| = 1\}$$

and positive real numbers $c_1, c_2, \dots, c_r$, where $1 \le r \le n$ in the real case and $1 \le r \le 2n-1$ in the complex case, such that

$$\sum_{i=1}^{r} c_i \overline{p^*(x_i)} p_0(x_i) = 0\,, \qquad p_0 \in P_0\,.$$

Since $L(p)p^* - L(p^*)p \in P_0$ for all $p \in P$, we have

$$L(p)\sum_{i=1}^{r} c_i |p^*(x_i)|^2 = L(p^*)\sum_{i=1}^{r} c_i \overline{p^*(x_i)} p(x_i)\,, \qquad p \in P\,.$$

Since $L(p^*) = \|L\|$ and $|p^*(x_i)| = 1$ for each i, we obtain (A.3.4) by taking

$$a_i = \frac{c_i \overline{p^*(x_i)}}{\sum_{j=1}^{r} c_j} \|L\|\,.$$

Using the fact that $|p^*(x_i)| = 1$ for each i in the above formula for a_i, we get (A.3.5). □

Comments, Exercises, and Examples.

We have followed Shapiro [71], which gives a long discussion of questions related to the best uniform approximation of a function $f \in C(Q)$ from a (usually finite-dimensional) linear subspace $P \subset C(Q)$. Some of these, together with other applications, are discussed in the exercises.

E.1 Caratheodory's Lemma. *If $A \subset \mathbb{R}^n$, then every point from the convex hull $co(A)$ of A can be written as a convex linear combination of at most $n+1$ points of A.*

Proof. Let $x \in A$. After a translation if necessary, we may assume that $x = 0$. Suppose

$$(A.3.6) \qquad 0 = \sum_{i=1}^{r} \alpha_i x_i\,, \qquad x_i \in A\,, \quad \alpha_i > 0\,, \quad r > n+1\,.$$

Since $r > n+1$, the elements $x_2, x_3, \dots, x_r$ are linearly dependent, so there exist real numbers β_i, $i = 2, 3, \dots, r$, not all zero, such that

$$\sum_{i=2}^{r} \beta_i x_i = 0\,.$$

Let $\beta_1 := 0$. For all $\lambda \in \mathbb{R}$, we have

$$0 = \sum_{i=1}^{r} \alpha_i x_i + \lambda \sum_{i=1}^{r} \beta_i x_i = \sum_{i=1}^{r} (\alpha_i + \lambda\beta_i) x_i \,.$$

When $\lambda = 0$, each term in the last sum is positive. We now define $c := \min |\alpha_i/\beta_i|$, where the minimum is taken for all indices i for which $\beta_i \neq 0$. If the index j is chosen so that $|\alpha_j/\beta_j| = c$, and if $\lambda := -\alpha_j/\beta_j$, then at least one of the numbers $\alpha_i + \lambda\beta_i$ is zero, and all are nonnegative. Also $\alpha_1 + \lambda\beta_1 = \alpha_1 > 0$. We have thus obtained a representation of the same form as (A.3.6), but with s terms, where $1 \leq s \leq r-1$. If $s > n+1$, then the process can be repeated, and after a finite number of steps we obtain the desired representation. □

E.2 Reformation in Terms of Integrals. Lemma A.3.2 can be reformulated as follows. Under the assumptions of Lemma A.3.2, $f \in C(Q)$ is orthogonal to P if and only if there exists a nonzero nonnegative Borel measure ρ on Q whose support consists of r points of $E(f)$ defined by (A.3.2), where $1 \leq r \leq n+1$ in the real case and $1 \leq r \leq 2n+1$ in the complex case, such that $\overline{f(x)}\, d\rho(x)$ annihilates P, that is,

$$\int_Q \overline{f(x)} p(x)\, d\rho(x) = 0\,, \qquad p \in P\,. \tag{A.3.7}$$

This reformation is not only a notational convenience, but it is essential in generalizations where P is no longer finite-dimensional. Moreover, (A.3.7) with any nonzero nonnegative Borel measure (not necessarily discrete) is sufficient for $f \perp P$. This is often useful, even when P is finite-dimensional; see Shapiro [71].

E.3 Haar's Characterization of Chebyshev Spaces.

a] Let $f_0, f_1, \ldots, f_n$ be real- or complex-valued continuous functions defined on a (not necessarily compact) Hausdorff space Q. Show that $P := \operatorname{span}\{f_0, \ldots, f_n\}$, where the span is taken over $\mathbb{R}$ (or $\mathbb{C}$), is a Chebyshev space if and only if there exists no real (or complex) measure on Q annihilating P whose support consists of less than $n+1$ points.

Hint: Use Proposition 3.1.2. □

b] Let P be an n-dimensional linear subspace of $C(Q)$, the space of real- (or complex-) valued continuous functions defined on a compact Hausdorff space Q containing at least n points. The space P is a Chebyshev space if and only if for each $f \in C(Q)$, there is a unique best uniform approximation to f from P.

Proof. First suppose P is a Chebyshev space of dimension n, and p_1 and p_2 are best uniform approximations to some $f \in C(Q)$ from P. Then $p_3 := \frac{1}{2}(p_1 + p_2)$ is also a best uniform approximation to f from P. As

$f - p_3 \perp P$, Lemma A.3.2 yields that $|f(x) - p_3(x)|$ attains its maximum on Q at r points, $x_1, x_2, \dots, x_r$, that support an annihilating measure for P, where $r \geq n+1$ by part a]. Note that

$$p_1(x_j) - f(x_j) = p_2(x_j) - f(x_j)\,, \qquad j = 1, 2, \dots, n+1\,,$$

and hence, since P is a Chebyshev space, $p_1 - p_2 = 0$.

Conversely, if P is not a Chebyshev space, then there exist n distinct points $x_1, x_2, \dots, x_n$ in Q such that the system of homogeneous linear equations

$$\begin{pmatrix} g_1(x_1) & \dots & g_1(x_n) \\ \vdots & \ddots & \vdots \\ g_n(x_1) & \dots & g_n(x_n) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\,,$$

where $\{g_1, \dots, g_n\}$ is a basis for P, has a nontrivial solution. Then also the homogeneous system formed with the transposed matrix has a nontrivial solution, so there exist constants b_i, not all zero, so that

$$\sum_{i=1}^{n} b_i g_i(x_j) = 0\,, \qquad j = 1, 2, \dots, n\,.$$

Thus, with $g := \sum_{i=1}^{n} b_i g_i$, we have

$$g(x_j) = 0\,, \qquad j = 1, 2, \dots, n\,.$$

We may assume, without loss, that $\|g\|_Q = 1$. Some of the constants $a_1, a_2, \dots, a_n$ may be zero; however, the set Γ of indices j for which $a_j \neq 0$ is not empty. By Tietze's theorem there exists an $f \in C(Q)$ such that $\|f\|_Q = 1$ and

$$f(x_j) = \frac{\overline{a}_j}{|a_j|}\,, \qquad j \in \Gamma\,.$$

Setting $h(x) := f(x)(1 - |g(x)|)$, we have

$$h(x_j) = \frac{\overline{a}_j}{|a_j|}\,, \qquad j \in \Gamma\,.$$

We claim that $\|h - p\|_Q \geq 1$ for every $p \in P$. Indeed, if $\|h - p\|_Q < 1$ for some $p \in P$, then

$$|f(x_j) - p(x_j)|^2 = |f(x_j)|^2 - 2\,\mathrm{Re}\big(p(x_j)\overline{f(x_j)}\big) + |p(x_j)|^2 < 1$$

for every $j \in \Gamma$; hence

$$\mathrm{Re}(a_j p(x_j)) > 0\,, \qquad j \in \Gamma\,.$$

Since $a_j = 0$ if $j \notin \Gamma$, we have $\mathrm{Re}(a_j p(x_j)) = 0$ for each $j \notin \Gamma$. Thus

$$\mathrm{Re}\left(\sum_{j=1}^{n} a_j p(x_j)\right) > 0\,.$$

However, if $p := \sum_{i=1}^{n} c_i g_i$, then

$$\sum_{j=1}^{n} a_j p(x_j) = \sum_{j=1}^{n} a_j \sum_{i=1}^{n} c_i g_i(x_j) = \sum_{i=1}^{n} b_i \sum_{j=1}^{n} a_i g_i(x_j) = 0\,,$$

which contradicts the previous inequality and shows that $\|h - p\|_Q \geq 1$ for every $p \in P$.

Finally, for all $\lambda \in [0,1]$, λg is a best uniform approximation to h from P because

$$\begin{aligned} |h(x) - \lambda g(x)| &\leq |f(x)|(1 - |g(x)|) + \lambda |g(x)| \\ &\leq 1 + (\lambda - 1)|g(x)| \leq 1 \end{aligned}$$

for all $x \in Q$, so the best uniform approximation to $h \in C(Q)$ from P is not unique. □

E.4 Unicity of the Extremal Function. Assume the notation of Theorem A.3.3. Show that if $P \subset C(Q)$ is an n-dimensional real Chebyshev space and $r = n$, then the extremal element $p^* \in P$ satisfying $\|p^*\|_Q = 1$ and $L(p^*) = \|L\|$ is unique.

The interesting relations of Theorem A.3.3 to the Riesz representation theorem, the Krein-Milman theorem, and the Hahn-Banach theorem are discussed in Shapiro [71].

E.5 Applications of the Interpolation Theorem. As before, let

$$D := \{z \in \mathbb{C} : |z| < 1\} \qquad \text{and} \qquad K := \mathbb{R} \pmod{2\pi}\,.$$

Prove the following statements. Each of them may be proven by characterizing the extremal polynomial for the given inequality with the help of Theorem A.3.3. A detailed hint is given only to part a].

a] Bernstein's Inequality.

$$|t'(\theta)| \leq n\|t\|_K\,, \qquad t \in \mathcal{T}_n\,, \quad \theta \in \mathbb{R}\,.$$

Hint: Let $\theta_0 \in \mathbb{R}$ be fixed, and study the linear functional $L(t) := t'(\theta_0)$, $t \in \mathcal{T}_n$. Observe that an extremal p in Lemma A.3.3 must satisfy $|p(x_i)| = 1$ for each $i = 1, 2, \ldots, r$ and r must equal $2n$. Note that $r \leq 2n$ holds by

Theorem A.3.3, while the argument for $r \geq 2n$ is similar to the corresponding step in the proof of Theorem 7.1.7. Finally, show that the extremal element t satisfying $L(t) = \|L\|$ is of the form

$$t(\theta) = \cos(n\theta - \alpha)$$

for some $\alpha \in K$. □

b] Markov's Inequality.

$$|p'(1)| \leq n^2 \|p\|_{[-1,1]}\,, \qquad p \in \mathcal{P}_n\,.$$

c] Chebyshev's Inequality.

$$|p(x)| \leq |T_n(x)|\,\|p\|_{[-1,1]}\,, \qquad p \in \mathcal{P}_n\,, \quad x \in \mathbb{R} \setminus [-1,1]\,,$$

where T_n is the Chebyshev polynomial of degree n defined by (2.1.1).

d] Bernstein's Inequality.

$$|p(z)| \leq |z|^n \|p\|_D\,, \qquad p \in \mathcal{P}_n^c\,, \quad z \in \mathbb{C} \setminus \overline{D}\,.$$

e] Bernstein's Inequality.

$$|p'(z)| \leq n|z|^{n-1} \|p\|_D\,, \qquad p \in \mathcal{P}_n^c\,, \quad z \in \mathbb{C} \setminus \overline{D}\,.$$

Hint: Use Theorem 1.3.1 (Lucas' theorem). □

f] Riesz' Identity. There are real numbers a_i with $\sum_{i=1}^{2n} |a_i| = n$ such that

$$t'(\theta) = \sum_{i=1}^{2n} a_i t(\theta + \theta_i)\,, \qquad t \in \mathcal{T}_n\,, \quad \theta \in \mathbb{R}\,,$$

where

$$\theta_i := \tfrac{2i-1}{2n}\,\pi\,, \qquad i = 1, 2, \ldots, 2n\,.$$

(This is, apart from the explicit determination of the number a_i, an identity discovered by M. Riesz [14].)

g] Show that in part f],

$$a_i = (-1)^{i+1} \frac{1}{4n \sin^2(\theta_i/2)}\,, \qquad i = 1, 2, \ldots, 2n\,.$$

h] Bernstein's Inequality in L_p.

$$\int_0^{2\pi} |t'(\theta)|^p \, d\theta \leq n^p \int_0^{2\pi} |t(\theta)|^p \, d\theta\,, \qquad t \in \mathcal{T}_n\,, \quad p \geq 1\,.$$

Hint: Use part f] and Jensen's inequality (see E.20 of Appendix 4). □

Arestov [81] shows that the inequality of part h] is valid for all $p > 0$. Golitschek and Lorentz [89] gives a simpler proof of this.

i] Find all extremal polynomials in parts a] to e] and h].

E.6 An Inequality of Szegő. If $p \in \mathcal{P}_n^c$ and $z_1, z_2, \ldots, z_{2n}$ are any equally spaced points on the unit circle ∂D, then

$$\|p'\|_D \leq n \max_{1 \leq k \leq 2n} |p(z_k)| .$$

Proof. See Frappier, Rahman, and Ruscheweyh [85]. □