第三章 函数逼近

朱升峰

华东师范大学 数学科学学院

2021 年 3 月

目录

基本概念

函数逼近是数值分析的基础。

插值是一种函数逼近。

本章讨论的函数逼近：$f(x) \in A$ (函数类)，在 B（函数类）中求 $p(x) \in B(\not\subseteq A)$，使 $p(x)$ 与 $f(x)$ 的误差在某种度量意义下最小。

- 线性相关、线性无关、基、坐标
- 有限维/无线维线性空间
- $\mathbb{R}^n$: n 维向量空间

  $C[a,b]$: $[a,b]$ 上连续函数集合

  $C^m[a,b]$: $[a,b]$ 上直到 m 阶导数都连续的函数集合

  H_n: 次数不超过 $(\leqslant) n$ 次的多项式集合

$C[a,b]$和$C^m[a,b]$是无线维的.

若函数 $p(x) \in H_n$,

$$p(x) = a_0 + a_1 x + \cdots a_n x^n, \quad \text{其中} a_i \text{为实数}$$

$p(x)$ 由系数唯一确定。$1, x, \cdots, x^n$ 线性无关，是 H_n 的一组基，
$H_n = \mathrm{span}\{1, x, \cdots, x^n\}$ 是 $n+1$ 维的，$(a_0, a_1, \cdots, a_n)$为 $p(x)$ 的坐标。

定理 1

设 $f(x) \in C[a, b]$，则对任何 $\epsilon > 0$，总存在一个代数多项式 $p(x)$，使

$$\max_{a \leqslant x \leqslant b} |f(x) - p(x)| < \epsilon$$

在 $[a, b]$ 上一致成立。

构造性证明

由于 $[a,b]$ 上函数可变量替换为 $[0,1]$ 上函数，不妨考虑 $[0,1]$ 上的 f。
Bernstein(1912) 给出证明，可参看《数值逼近》(王仁宏著)。
构造多项式

$$B_n(f,x)=\sum_{k=0}^{n} f\left(\frac{k}{n}\right) P_k(x),$$

其中 $P_k(x)=\binom{n}{k}x^k(1-x)^{n-k}$ 和 $\binom{n}{k}=\frac{n(n-1)\cdots(n-k+1)}{k!}$
Bernstein 证明了在 $[0,1]$ 上一致成立：

$$\lim_{n\to\infty} B_n(f,x)=f(x)$$

若 $f(x)$ 在 [0,1] 上 m 阶导数连续，则

$$\lim_{n\to\infty} B_n^{(m)}(f,x)=f^{(m)}(x)$$

函数逼近问题

$\varphi(x) \in \Phi \ = \text{span}\{\varphi_i(x)\}_{i=0}^n$ 可表示为

$$\varphi(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_n\varphi_n(x).$$

其中 $\{\varphi_i(x)\}_{i=0}^n \subset C[a,b]$ 线性无关。

函数逼近问题：对任一 $f \in C[a,b]$，在子空间 Φ 中找一个元素（函数）$\varphi^*(x) \in \Phi$，使 $f(x) - \varphi^*(x)$ 在某种度量意义下最小。

定义 2

设 S 为线性空间，$x \in S$，若存在唯一实数 $\|\cdot\|$ 满足条件：

- **正定性**：$\|x\| \geqslant 0$，当且仅当 $x = 0$ 时，$\|x\| = 0$.
- **齐次性**：$\|\alpha x\| = |a|\|x\|, \alpha \in \mathbb{R}$.
- **三角不等式**：$\|x + y\| \leqslant \|x\| + \|y\|$.

则称 $\|\cdot\|$ 为线性空间 S 上的范数，S 称为赋范线性空间。

赋范线性空间

$\mathbb{R}^n$: $x=(x_1,x_2,\cdots,x_n)^T\in\mathbb{R}^n$，可定义三种常用范数

- ∞-范数或最大范数

$$\|x\|_\infty=\max_{1\leqslant i\leqslant n}|x_i|$$

- 1-范数

$$\|x\|_1=\sum_{i=1}^n|x_i|$$

- 2-范数

$$\|x\|_2=\Big(\sum_{i=1}^n x_i^2\Big)^{1/2}$$

赋范线性空间

$C[a,b]$: $f\in C[a,b]$，可定义三种常用范数

- ∞-范数或最大范数

$$\|f\|_\infty = \max_{a\leqslant x\leqslant b} |f(x)|$$

- 1-范数

$$\|f\|_1 = \int_\Omega |f(x)|\mathrm{d}x$$

- 2-范数

$$\|f\|_2 = \left(\int_\Omega f(x)^2\mathrm{d}x\right)^{1/2}$$

定义 3

设 X 是数域 K($\mathbb{R}$ 或 $\mathbb{C}$) 上的线性空间，$\forall u, v \in X$，有 K 中一个数与之对应，记为 (u, v)，满足条件：

(1) $(u, v) = \overline{(v, u)}, \forall u, v \in X$;

(2) $(\alpha u, v) = \alpha(u, v), \alpha \in K, u, v \in X$;

(3) $(u + v, w) = (u, w) + (v, w), \forall u, v, w \in X$

(4) $(u, u) \geqslant 0$, 当且仅当 $u = 0$ 时 , $(u, u) = 0$.

则称 (u, v) 为 X 上 u 与 v 的内积。

内积空间：定义了内积的线性空间。当 $K = \mathbb{R}$，(1) 即 $(u, v) = (v, u)$

如果 $(u, v) = 0$，则称 u 与 v正交，这是向量垂直概念的推广。

Cauchy-Schwarz 不等式

定理 4

设 X 为内积空间，对 $\forall u, v \in X$，有

$$|(u,v)|^2 \leqslant (u,u)(v,v).$$

证明.

当 $v=0$ 时，显然成立。设 $v \neq 0$，对任何数 λ 有
$0 \leqslant (u+\lambda v, u+\lambda v) = (u,u) + 2\lambda(u,v) + \lambda^2(v,v)$. 取 $\lambda = -(u,v)/(v,v)$, 代人右端, 得

$$(u,u) - 2\frac{|(u,v)|^2}{(v,v)} + \frac{|(u,v)|^2}{(v,v)} \geqslant 0,$$

由此即得 $v \neq 0$ 时

$$|(u,v)|^2 \leqslant (u,u)(v,v).$$

□

定理 5

设 X 为内积空间，$u_1, u_2, \cdots, u_n \in X$，Gram 矩阵

$$G = \begin{pmatrix} (u_1, u_1) & (u_2, u_1) & \cdots & (u_n, u_1) \\ (u_1, u_2) & (u_2, u_2) & \cdots & (u_n, u_2) \\ \vdots & \vdots & & \vdots \\ (u_1, u_n) & (u_2, u_n) & \cdots & (u_n, u_n) \end{pmatrix}$$

非奇异的充分必要条件是 $u_1, u_2, \cdots, u_n$ 线性无关。

证明：G 非奇异等价于行列式 $\det G \neq 0$，等价于关于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性方程组

$$\left(\sum_{j=1}^{n} \alpha_j u_j, u_k\right) = \sum_{j=1}^{n} (\alpha_j u_j, u_k) = 0, \ \ k = 1, 2, \cdots, n$$

只有零解。

注意到上式

$$\left(\sum_{j=1}^{n}\alpha_j u_j, u_k\right) = 0, \quad k = 1, 2, \cdots, n$$

$$\Leftrightarrow \left(\sum_{j=1}^{n}\alpha_j u_j, \sum_{j=1}^{n}\alpha_j u_j\right) = 0$$

$$\Leftrightarrow \sum_{j=1}^{n}\alpha_j u_j = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0.$$

线性无关意味着上式只有零解 $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$，证毕。
内积空间上可以由内积导出一种范数，即

$$\|u\| = \sqrt{(u, u)}, \quad u \in X$$

容易验证满足范数定义的三个条件。

例: $\mathbb{R}^n$ 的内积

设 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$， $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T, \boldsymbol{y} = (y_1, y_2, \cdots, y_n)^T$, 则其内积定义为

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} x_i y_i.$$

由此导出向量 2-范数

$$\|\boldsymbol{x}\|_2 = (\boldsymbol{x}, \boldsymbol{x})^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}.$$

一般地，若给定实数 $\omega_i > 0 (i = 1, 2, \cdots, n)$，称之为权系数，则可以定义加权内积

$$\|\boldsymbol{x}\|_2 = \left(\sum_{i=1}^{n} \omega_i x_i^2\right)^{1/2}.$$

内积

例: $C[a,b]$ 的内积

设函数 $f,g\in C[a,b]$，则其内积定义为

$$(f,g)=\int_a^b fg\mathrm{d}x.$$

由此导出 2-范数

$$\|f\|_2=(f,f)^{1/2}=\left(\int_a^b f(x)^2\mathrm{d}x\right)^{1/2}.$$

定义 6

$[a,b]$ 是有限或无限区间，$[a,b]$ 上的非负函数 $\rho(x)$ 满足条件：

(1) $\int_a^b x^k\rho(x)\mathrm{d}x$ 存在且为有限值 $(k=0,1,\cdots)$;

(2) 对 $[a,b]$ 上非负连续函数 $g(x)$，若 $\int_a^b \rho(x)g(x)\mathrm{d}x$ 则 $g(x)\equiv 0$,

则称 $\rho(x)$ 为 $[a,b]$ 上的权函数。

带权内积

设函数 $f(x), g(x) \in C[a,b]$，$\rho(x)$ 是 $[a,b]$ 上的权函数，则带权 $\rho(x)$ 的内积定义为

$$(f,g) = \int_a^b \rho(x) f(x) g(x) \mathrm{d}x.$$

由此导出的带权 $\rho(x)$ 的范数

$$\|f\|_2 = (f,f)^{1/2} = \left(\int_a^b \rho(x) f(x)^2 \mathrm{d}x \right)^{1/2}.$$

最佳逼近

设 $\varphi_0, \varphi_1, \cdots, \varphi_n$ 线性无关，记 $\Phi = \mathrm{span}\{\varphi_0, \varphi_1, \cdots, \varphi_n\}$，Gram 矩阵为

$$G = G(\varphi_0, \varphi_1, \cdots, \varphi_n) = \begin{bmatrix} (\varphi_0, \varphi_0) & (\varphi_0, \varphi_1) & \cdots & (\varphi_0, \varphi_n) \\ (\varphi_1, \varphi_0) & (\varphi_1, \varphi_1) & \cdots & (\varphi_1, \varphi_n) \\ \vdots & \vdots & & \vdots \\ (\varphi_n, \varphi_0) & (\varphi_n, \varphi_1) & \cdots & (\varphi_n, \varphi_n) \end{bmatrix}.$$

函数逼近：对给定 $f \in C[a, b]$，寻求它的最佳逼近多项式。

- 若 $P^*(x) \in H_n$，误差范数

$$\|f(x) - P^*(x)\| = \min_{P \in H_n} \|f(x) - P(x)\|,$$

则称 $P^*(x)$ 是 $[a, b]$ 上的最佳逼近多项式。

- 如果 $P^*(x) \in \Phi = \mathrm{span}\{\varphi_0, \varphi_1, \cdots, \varphi_n\}$，相应的称 $P^*(x)$ 为最佳逼近函数。

常用的最佳逼近

- $\|\cdot\|_\infty$

$$\begin{aligned}\|f(x)-P^*(x)\|_\infty &= \min_{P\in H_n}\|f(x)-P(x)\|_\infty\\ &= \min_{P\in H_n}\max_{a\leqslant x\leqslant b}|f(x)-P(x)|\end{aligned}$$

称 $P^*(x)$ 为最佳一致逼近多项式。

- $\|\cdot\|_2$

$$\begin{aligned}\|f(x)-P^*(x)\|_2^2 &= \min_{P\in H_n}\|f(x)-P(x)\|_2^2\\ &= \min_{p\in H_n}\int_a^b [f(x)-P(x)]^2\,\mathrm{d}x\end{aligned}$$

称 $P^*(x)$ 为最佳平方逼近多项式。

最小二乘拟合

已知节点上给出函数值 $y_i := f(x_i)\,(i = 0, 1, \cdots, m)$，数据表

x	x_0	x_1	$\dots$	x_m
$f(x)$	y_0	y_1	$\dots$	y_m

找 $g^*(x) \in \Phi$, 使得

$$\sum_{i=0}^{m} \left|y_i - g^*(x_i)\right|^2 = \min_{g\in\Phi} \sum_{i=0}^{m} \left|y_i - g(x_i)\right|^2$$

称 $g^*(x)$ 为 $f(x)$ 的最小二乘拟合函数。

若 $\Phi = H_n (m > n)$，则称 $g^*(x)$ 为 n 次最小二乘拟合多项式。

最小二乘是数据/曲线拟合方法中常用的方法。