多项式最佳逼近

假定 $\mathcal{H}$ 是 $C[a,b]$ 中的一个线性子空间，对任给的 $f \in C[a,b]$ ，我们寻找一 $p^* \in \mathcal{H}$ 使得 p^* 是 $\mathcal{H}$ 中对 f 的最佳逼近。我们通常选择 $\mathcal{H}$ 为次数不超过 n 的多项式空间，记为 $\mathbb{P}_n$ 。当然，这一“最佳”也依赖于我们所选的度量方式——范数。我们可更精确的将该问题描述为：对于给定的函数 $f \in C[a,b]$ 和范数 X ，寻找一次数不超过 n 的多项式 p^* ，使得 p^* 是在范数 X 意义下对 f 的最佳逼近，即：

$$\|p^* - f\|_X = \min_{p \in \mathbb{P}_n} \|p - f\|_X.$$

当然，选择不同的范数将导致最终结果不同。我们在此考虑三种不同的范数。

1. L^∞ 范数

如果我们选择范数 X 为 L_∞ 范数，那么我们事实上是求解如下问题：

$$\inf_{p \in \mathbb{P}_n} \max_{a \le x \le b} |p(x) - f(x)|.$$

我们将上述问题的解记为 p^*, 并令

$$e(x) := p^*(x) - f(x),$$

及

$$\mathcal{A} := \{x \in [a,b] : |f(x) - p^*(x)| = \max_{a \le x \le b} |e(x)|\}.$$

那么，下面的 Kolmogorov 最佳逼近定理给了一个 p^* 的刻画：

定理 1. 多项式 $p^* \in \mathbb{P}_n$ 是函数 $f \in C[a,b]$ 在 $\mathbb{P}_n$ 中的最佳逼近多项式，必须且只须不存在 $q \in \mathbb{P}_n$ 使得

$$(f(x) - p(x))q(x) > 0, \qquad \forall x \in \mathcal{A}.$$

下面的 Chebyshev 定理给出了更为具体的刻画：

定理 2. 函数 f 在 $\mathbb{P}_n$ 中的最佳逼近多项式是存在且唯一的。多项式 $p \in \mathbb{P}_n$ 是 f 的最佳逼近多项式必须且只须 $f - p$ 在 $[a,b]$ 上点数不少于 $n+2$ 的点列

$$x_1 < x_2 < \cdots < x_N \qquad (N \ge n+2)$$

上以正负交错的符号取到 $\max_{a\le x\le b}|f(x)-p(x)|$ 的值。

作为 Chebyshev 定理的一个应用，我们可以证明 Chebyshev 多项式在区间 $[-1,1]$ 的一种最小性质。回忆 Chebyshev 多项式定义为

$$T_n(x)=\cos(n\arccos x).$$

一个简单的观察是 T_n 满足如下的递归关系

$$T_{n+1}(x)=2xT_n(x)-T_{n-1}(x),\quad n=1,2,\ldots,$$

且 $T_n/2^{n-1}$ 为首项为 x^n 的多项式。

定理 3. 假定 $x\in[-1,1]$ 且令 n 为任意正整数.那么对任意的首项为 x^n 的多项式 $p\in\mathbb{P}_n$，我们有

$$\max_{-1\le x\le 1}T_n(x)/2^{n-1}\le\max_{-1\le x\le 1}p(x).$$

2. L^1 范数

如果我们选取 X 为 L^1 范数，那么我们事实上是考虑如下问题：

$$\inf_{p\in\mathbb{P}_n}\int_a^b|f(x)-p(x)|dx.$$

L^1 最佳逼近亦是逼近论中的经典内容，但过去对其重视不足。近些年，随着稀疏逼近，特别是压缩感知的兴起，人们从而对 L^1 最佳逼近有了更多的新认识。

我们假定 $p^*\in\mathbb{P}_n$ 是对 f 的最佳 L^1 逼近。我们定义如下的符号函数：

$$s^*(x)=\begin{cases}-1, & f(x)<p^*(x)\\ 0, & f(x)=p^*(x)\\ 1, & f(x)>p^*(x),\end{cases}\qquad a\le x\le b.$$

下面的定理刻画了 L^1 最佳逼近多项式的特征。

定理 4. 令 $f\in C[a,b]$。令 $f\in\mathbb{P}_n$ 且使得集合

$$\mathcal{E}:=\{x:f(x)=p^*(x),a\le x\le b\}$$

为空或者由有限个点或区间组成。那么 p^* 是 $\mathbb{P}_n$ 中在范数 L^1 下对 f 的最佳逼近，如果和仅仅如果不等式

$$\left|\int_a^b s^*(x)p(x)dx\right|\le\int_{\mathcal{E}}|p(x)|dx$$

对所有 $p\in\mathbb{P}_n$ 均成立。

很多情况下，集合 $\mathcal{E}$ 是一个离散集合，那么 $\int_{\mathcal{E}}|p(x)|dx=0$. 那么， p^* 是对 f 的 L^1 最佳逼近，如果和仅仅如果

$$\int_a^b s^*(x)p(x)=0, \quad \forall p\in\mathbb{P}_n.$$

根据该定理，我们很容易得到，单调函数 f 在 $\mathbb{P}_0$ 中的最佳逼近为 $p^*(x)=f((a+b)/2)$.

在实际中，特别是信号处理中，人们可能仅仅知道函数 f 在一些离散点上得值。因此，人们需要考虑离散情形的 ℓ^1 最佳逼近。我们在此考虑加权 ℓ^1 范数：即对向量 $f\in\mathbb{R}^m$, 和向量 $w=(w_1,\dots,w_m), w_i>0$, 那么

$$\|f\|_{w,1}:=\sum_{i=1}^m w_i|f_i|.$$

所谓离散 ℓ^1 最佳逼近，即指对一线性子空间 $\mathcal{H}\subset\mathbb{R}^m$ ，寻找 $p^*\in\mathcal{H}$ 使得

$$\|p^*-f\|_{w,1}=\min_{p\in\mathcal{H}}\|p-f\|_{w,1}.$$

我们将导出离散 ℓ^1 最佳逼近的特征。对于 $p^*\in\mathcal{H}$ ，我们定义符号向量：

$$s^*(t)=\begin{cases}-1, & f(t)<p^*(t)\\ 0, & f(t)=p^*(t)\\ 1, & f(t)>p^*(t),\end{cases} \qquad t=1,\dots,m.$$

定理 5. 向量 $p^*\in\mathcal{H}$ 是 $f\in\mathbb{R}^m$ 的 ℓ^1 最佳逼近如果和仅仅如果下述不等式对所有 $p\in\mathcal{H}$ 均成立：

$$\left|\sum_{t=1}^m w(t)s^*(t)p(t)\right|\le\sum_{t\in\mathcal{Z}}w(t)|p(t)|,$$

此处

$$\mathcal{Z}:=\{t:p^*(t)=f(t), 1\le t\le m\}.$$

更进一步，我们有

定理 6. 假定 $\dim\mathcal{H}=n$. 假定向量 $p^*\in\mathcal{H}$ 是 $f\in\mathbb{R}^m$ 的唯一 ℓ^1 最佳逼近，并令

$$\mathcal{Z}:=\{t:p^*(t)=f(t), 1\le t\le m\}.$$

那么,

$$\#\mathcal{Z}\ge n.$$

我们下面简单介绍离散最佳 ℓ^1 逼近与压缩感知之间的关联。假定信号 $f\in\mathbb{R}^m$ 是一稀疏信号，即其非零元素很少。我们将 f 中非零元素数目记为 k ，且 $k\ll m$. 那么， $\mathcal{H}$ 中对 f 的 ℓ^1

最佳逼近是 $\mathcal{H}$ 中的 0 元素的一个充分条件是：对任意的 $T \subset \{1, \ldots, m\}$ 且 $\#T \le n-k$ ，我们有

$$\sum_{t \notin T} w(t)|p(t)| \le \sum_{t \in T} w(t)|p(t)|, \tag{1}$$

对任意的 $p \in \mathcal{H}$ 均成立。条件 (1) 也被称为 **零空间性质** 。实际中，人们通常选择 $\mathcal{H}$ 为一线性方程组的零解空间，即存在一 $(m-n) \times m$ 的矩阵 A 使得

$$\mathcal{H} = \{x \in \mathbb{R}^m : Ax = 0\}.$$

那么，一个自然的问题是：矩阵 A 满足什么样的条件才能保证 $\mathcal{H}$ 满足零空间性质呢？为此，人们引入了 RIP 条件。我们称矩阵满足 k 阶 RIP 条件，即对任意的 $T \subset \{1, \ldots, m\}$ 且 $\#T \le k$ 存在常数 δ 使得：

$$(1-\delta)\|x_T\|^2 \le \|\Phi_T x_T\|^2 \le (1+\delta)\|x_T\|^2$$

成立。关于 RIP 条件与零空间性质是当前压缩感知中的一个热门研究领域。这方面的更多介绍可参考 [2, 1].

3. L^2 范数

设 ρ 是一个在区间 $[a,b]$ 上 L 可积的非负函数，它至多只在一个测度为 0 的集合上可能等于 0 。我们称 ρ 为一个权函数。对于任意一个定义在区间 $[a,b]$ 上的可测函数 f ，如果 $\rho \cdot f$ 是 L 可积的，则说 f 属于 $L_\rho[a,b]$ 类，如果 $\rho \cdot f^2$ 是 L 可积的，则说 f 属于 $L^2_\rho[a,b]$ 类。$L^2_\rho[a,b]$ 中的每一个函数 f ，其范数定义为

$$\|f\|_{\rho,2} = \sqrt{\int_a^b \rho(x) f^2(x) dx}.$$

那么，$\|f-g\|_{\rho,2}$ 给出两个函数 f, g 之间的距离。所谓 L^2 最佳逼近就是对任意的函数 $f \in C[a,b]$ 寻找一函数 $p \in \mathbb{P}_n$ 使得

$$\|f-p\|_{\rho,2}$$

达到最小。

设 ρ 为定义在区间 $[a,b]$ 上的权函数。如果函数 f 与 g 满足条件：

$$\int_a^b \rho(x) f(x) g(x) dx = 0$$

则说 f 和 g 在 $[a,b]$ 上关于权函数 ρ 是直交的。如果函数系统

$$\omega_1, \omega_2, \ldots$$

中的每一对函数在区间 $[a,b]$ 上关于权函数 ρ 均直交，则称该系统为 $[a,b]$ 上关于权函数 ρ 的直交函数系。我们看几个例子：

例 Legender 多项式系

$$p_n(x) = \frac{1}{2^n n!}(\frac{d}{dx})^n (x^2-1)^n, \quad n = 0, 1, 2, 3, \ldots$$

是区间 $[-1,1]$ 上的直交多项式。

例 Tchebyshev 多项式系

$$T_n(x) = \cos(n \arccos x), \quad n = 0, 1, 2, \ldots$$

是区间 $[-1,1]$ 上对权函数 $(1-x^2)^{-1/2}$ 的直交系。

下面我们介绍广义 Fourier 展开的问题。设置

$$A_k = \int_a^b \rho(x)\omega_k^2(x)dx.$$

和

$$c_k = \frac{1}{A_k}\int_a^b \rho(x)\omega_k(x)f(x)dx, k = 1, 2, \ldots$$

从而有如下的广义 Fourier 级数：

$$f(x) \ \sim \ \sum_{k=1}^{\infty} c_k \omega_k(x)$$

定理 7. 对于任意给定的正整数 n ，用线性组合式

$$F(x) = \sum_{k=1}^{n} a_k \omega_k(x)$$

做成的函数对 f 进行平方逼近，为使偏差

$$\|F-f\|_{\rho,2} = (\int_a^b \rho(x)(F(x)-f(x))^2 dx)^{1/2}$$

达到最小，函数 F 必须等于广义 *Fourier* 级数的部分和 $S_n(x) := \sum_{k=1}^{n} c_k \omega_k(x)$ 。而偏差最小值等于

$$\left(\int_a^b \rho(x)f^2(x)dx - \sum_{k=1}^{n} A_k c_k^2\right)^{1/2}.$$

注意到 $\|S_n - f\| \geq 0$ ，因此，可有

$$\sum_{k=1}^{n} A_k c_k^2 \ \leq \ \int_a^b \rho(x) f^2(x) dx.$$

令 $n \to \infty$ 有

$$\sum_{k=1}^{\infty} A_k c_k^2 \leq \int_a^b \rho(x) f^2(x) dx,$$

这就是 Bessel 不等式。根据偏差的最小值表达式可知，上述 Bessel 不等式能改为 Parseval 等式

$$\sum_{k=1}^{\infty} A_k c_k^2 \ = \ \int_a^b \rho(x) f^2(x) dx$$

的充要条件是

$$\lim_{n\to\infty} \|S_n - f\|_{\rho,2} = 0.$$

我们简要介绍三种最常用的直交多项式，它们是 Legender 多项式系 $\{p_n\}$ ， Laguerre 多项式系 $\{L_n\}$ 和 Hermite 多项式系 $\{H_n\}$ 。它们在数学物理问题及数值积分中均有重要意义。

1. Legender 多项式系

在区间 $[-1,1]$ 上对于权函数 $\rho = 1$ 构成的直交系的多项式 p_n 称为 Legendre 多项式系，其简单表达形式为

$$p_n(x) = \frac{1}{2^n n!} (\frac{d}{dx})^n (x^2 - 1)^n.$$

$p_n(x)$ 实际上是下列 Legendre 微分方程式

$$\frac{d^2 y}{dx^2} - \frac{2x}{1-x^2}\frac{dy}{dx} + \frac{n(n+1)}{1-x^2} y = 0$$

在正规点 $x = 0$ 附近满足条件 $y(1) = 1$ 的唯一确定的多项式解。 $p_n(x)$ 有如下的母函数

$$\frac{1}{\sqrt{1 - 2xz + z^2}} = \sum_{n=1}^{\infty} p_n(x) z^n, \quad |z| < 1.$$

2. Laguerre 多项式系

以前我们所讨论的一切，都是假定基本区间 $[a,b]$ 是有限的。其实，权函数、 L^2_ρ 空间以及直交系等概念也完全可以推广到无限区间的情形。所谓的 Laguerre 多项式系 $\{L_n\}$ ，就是在区间 $(0, +\infty)$ 上关于权函数 e^{-x} 所构成的直交系。它们可以表达为：

$$L_n(x) = e^x \left(\frac{d}{dx}\right)^n (x^n e^{-x}).$$

只要将上式右端的导数算出，就知道 L_n 是 n 次多项式

$$L_n(x) = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} n(n-1)\cdots(n-k+1)x^{n-k}.$$

3. Hermite 多项式系

所谓的 Hermite 多项式 $\{H_n\}$ 就是在区间 $(-\infty, \infty)$ 上关于权函数 e^{-x^2} 所构成的直交系。它可以通过如下的表达式来定义：

$$H_n(x) = e^{x^2} \left(\frac{d}{dx}\right)^n (e^{-x^2}).$$

References

[1] A. Cohen, W. Dahmen and R. DeVore, Compressed sensing and best k-term approximation, J. Amer. Math. Soc., 22 (2009)211-231.

[2] E. J. Candès, J. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm. Pure Appl. Math., 59(8)(2006)1207-1223.