从数据拟合到积分变换

From Data Fitting to Integral Transforms

纪浩正 / Haozheng Ji
Department of Earth and Space Sciences, Southern University of Science and Technology
12311405@mail.sustech.edu.cn

March 28, 2025

Abstract

本文系统探讨了数据拟合与积分变换的理论框架及方法应用。研究以函数空间正交投影为核心，阐释了Gram-Schmidt正交化在构建基函数中的作用，通过多项式、傅里叶基、指数函数等典型基的对比实验，揭示了基函数选择对拟合效果的影响机制。进一步推导了傅里叶变换的正交投影本质，拓展至拉普拉斯变换的积分表达式，到各类积分变换，建立了微分方程求解的新视角。

Keywords: 数据拟合；Gram-Schmidt正交化；傅里叶变换；拉普拉斯变换；积分变换；函数空间投影

1 拟合

基本定义：
拟合指的是寻找一个函数 $f(x)$ 来近似描述给定的数据集 $\{x_i, y_i\}_{i\in\mathbf{N}}$，使得每个 $f(x_i)$ 都尽可能接近对应的 y_i。衡量拟合效果的一个常用指标是残差平方和

$$\sum_{i=1}^{N} |f(x_i) - y_i|^2.$$

类似地，若要比较两个函数 $f(x)$ 与 $g(x)$ 之间的差异，也可以采用类似的标准，其连续形式可写为

$$\int_a^b |f(x) - g(x)|^2 \, dx.$$

在函数向量空间中，我们定义内积为

$$\langle f, g\rangle = \int_a^b f(x)g(x)\, dx,$$

进而定义范数

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

从而 $\|f - g\|$ 自然地代表了 f 与 g 之间的距离。

如何进行拟合

在高中阶段，我们学习了利用最小二乘法拟合数据点，但该方法仅适用于线性关系，而自然界中的线性现象往往较少。

最小二乘法利用形式如

$$y = \hat{a} \cdot 1 + \hat{b} \cdot x,$$

的方程，将 $\{1, x\}$ 作为基来拟合数据。随后，通过泰勒展开，我们开始使用 $\{1, x, x^2, x^3, \ldots\}$ 作为基函数来逼近函数。进一步，我们又引入了傅里叶级数

$$\{1, \cos x, \cos 2x, \cos 3x, \ldots, \sin x, \sin 2x, \sin 3x, \ldots\}$$

作为拟合工具。

以 $f(x) = 3x^2 + 2x - 1$ 为例，若采用最小二乘法，只能在有限区间内捕捉其大致变化趋势；而若采用泰勒级数，仅用前三项便可完全表达 $f(x)$ 所蕴含的信息，实现反演与预测时都能获得精确结果。反之，若采用傅里叶级数，在给定闭区间内

$$\sum_{i=-n}^{n} C_i e^{j\omega_i x}$$

可以绝对收敛于 $f(x) = 3x^2 + 2x - 1$；但当超出该区间时，由于傅里叶级数由正弦与余弦函数构成，表现出明显周期性，而 $f(x)$ 则无此周期性，因此拟合将失效。这正是所选拟合基不同所带来的差异。虽然我们无法总是选出完美的基函数进行预测，但在给定范围内使用适当的基函数可以有效反映数据或函数的变化。

2 Gram–Schmidt 正交化

函数拟合方法

我们利用基函数集 $\{v_1(x), v_2(x), v_3(x), \ldots, v_n(x)\}$ 来拟合目标函数。该集合张成的函数空间 $\mathbf{V}$ 中的任一函数都可表示为

$$u(x) = \sum_{i=1}^{n} C_i v_i(x).$$

我们的目标是确定一个 $u(x)$，使得残差平方和

$$\sum_{i=1}^{N} |f(x_i) - u(x_i)|^2$$

最小。也可引入权函数 $r(x)$ 考虑不同数据点的重要性，此时残差平方和变为

$$\sum_{i=1}^{N} |f(x_i) - u(x_i)|^2 r(x)$$

或其连续形式

$$\int_a^b |f(x) - u(x)|^2 r(x)\, dx.$$

为简便起见，我们通常取 $r(x) = 1$。因此，表达式

$$\int_a^b |f(x) - u(x)|^2\, dx$$

反映了 f 与 u 在整体空间 $\mathbf{W}$ 中的距离。要使

$$\int_a^b |f(x)-u(x)|^2\,dx$$

最小，则 u 必须为 f 在空间 $\mathbf{V}$ 上的正交投影。

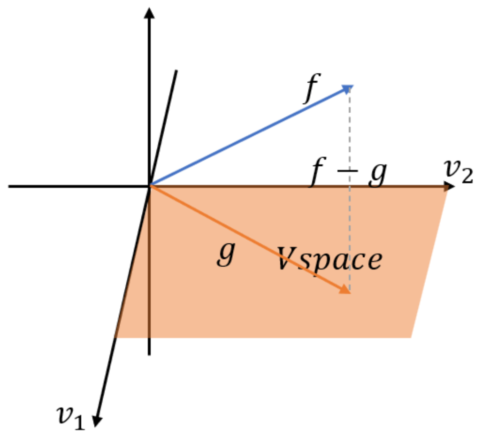

Figure 1: 正交投影示意图

构造正交归一基
在具有内积定义的向量空间中，我们可采用 Gram–Schmidt 正交化过程将基 $\{v_1(x), v_2(x), v_3(x), \dots, v_n(x)\}$ 转化为正交归一基 $\{u_1(x), u_2(x), u_3(x), \dots, u_n(x)\}$：

$$\begin{aligned}
&\textbf{Step 1:}\quad u_1(x)=\frac{v_1(x)}{\sqrt{\langle v_1,v_1\rangle}}=\frac{v_1(x)}{\sqrt{\int_a^b v_1(x)^2\,dx}},\\
&\textbf{Step 2:}\quad u_2'(x)=v_2(x)-\langle v_2,u_1\rangle\,u_1(x),\quad u_2(x)=\frac{u_2'(x)}{\sqrt{\langle u_2',u_2'\rangle}},\\
&\qquad\vdots\\
&\textbf{Step n:}\quad u_n'(x)=v_n(x)-\sum_{i=1}^{n-1}\langle v_n,u_i\rangle\,u_i(x),\quad u_n(x)=\frac{u_n'(x)}{\sqrt{\langle u_n',u_n'\rangle}}.
\end{aligned}$$

不难验证，$\langle u_i, u_j\rangle=\delta_{ij}$。例如：

$$\begin{aligned}
\textbf{例子:}\quad \langle u,v\rangle &= \int_{-1}^{1} u(x)v(x)\,dx,\\
\{1,x,x^2\} &\Rightarrow \left\{\frac{1}{\sqrt{2}},\ \frac{\sqrt{3}\,x}{\sqrt{2}},\ \frac{3\sqrt{10}}{4}x^2-\frac{\sqrt{10}}{4}\right\},\\
\{1,\cos\pi x,\cos 2\pi x\} &\Rightarrow \left\{\frac{1}{\sqrt{2}},\ \cos\pi x,\ \cos 2\pi x\right\}.
\end{aligned}$$

获得投影
有了正交归一基后，我们便可将 f 投影到 $\mathbf{V}$ 上，其投影记为 $P_V f$：

$$P_V f=\sum_{i=1}^{n}\langle f,u_i\rangle\,u_i.$$

这样，f 与空间 $\mathbf{V}$ 之间的距离（即残差平方和 $\int_a^b |f(x) - P_V f(x)|^2 \, dx$）便取得最小值。

例如，我们使用基 $\{\sin \pi x, \sin 2\pi x, \sin 3\pi x, \ldots \sin n\pi x$ 来拟合函数

$$u(x) = \begin{cases} 0, & -1 < x < 0, \\ 1, & 0 \leq x < 1, \end{cases}$$

可以清楚地看到，这正是傅里叶级数展开的过程；由于各项正交，每次新增的项均展现出正弦函数的形状。

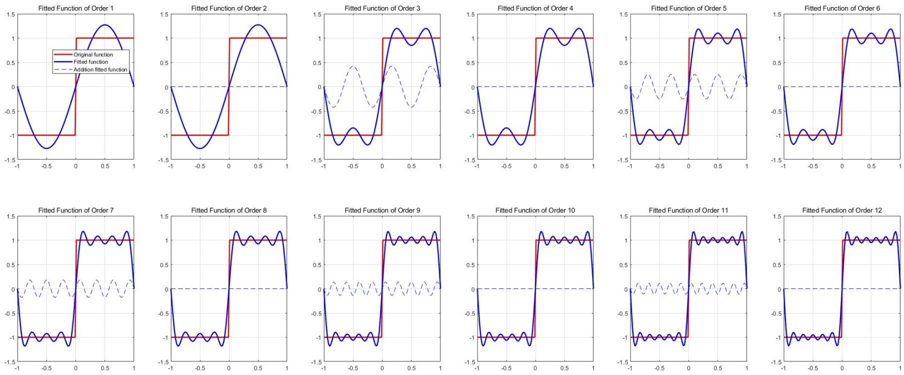

Figure 2: 正弦级数拟合示意图

下图展示了使用基

$$\left\{ \sin\!\Big(\Big(\frac{1 - (-1)^n + 2n}{4} \Big) \pi x + \frac{\pi\,(1 + (-1)^n)}{4} \Big) \right\}_{0 \leq n \leq N}$$

（即余弦与正弦交替构成的基）进行展开的效果。可以看出，在余弦项上，附加项均为零，这是因为余弦函数为偶函数，而待拟合的方波为奇函数，在对称区间上它们天然正交，因此投影为零。由此可见，余弦函数不能反映出方波函数的特性，我们需要补充一组奇函数基来拟合目标函数。

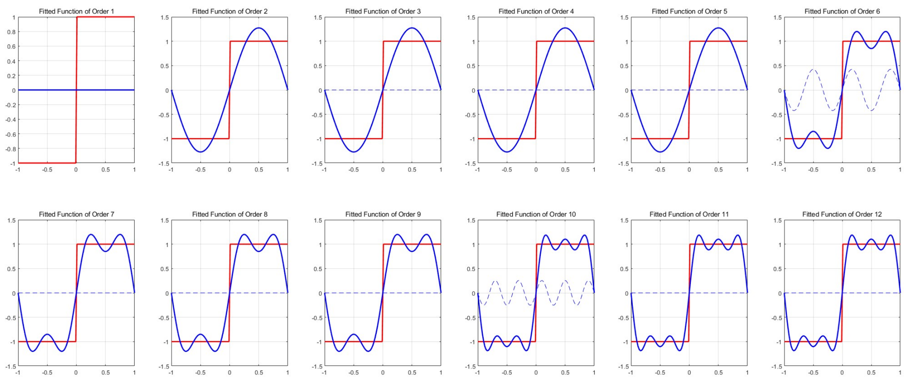

Figure 3: 交替余弦与正弦基的傅里叶展开

同理，使用基 $\{1, x, x^2, x^3, \ldots x^n\}$ （需先进行正交归一化）拟合时，前几项表现出明显的多项式函数的特性，而后续项则因是多个多项式函数的线性组合而表现出模糊的多项式函数特性。

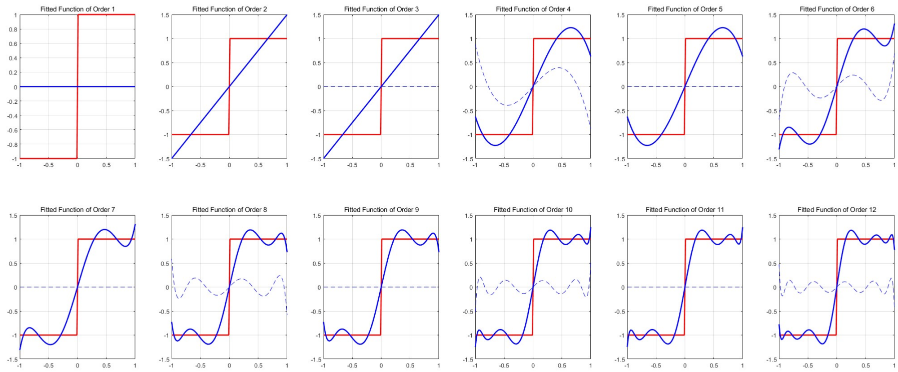

Figure 4: 多项式基拟合示意图

此外，我们还可以采用其他函数进行拟合，例如基 $\{1, e^x, e^{2x}, e^{3x}, \dots e^{nx}\}$（需先进行正交归一化）。

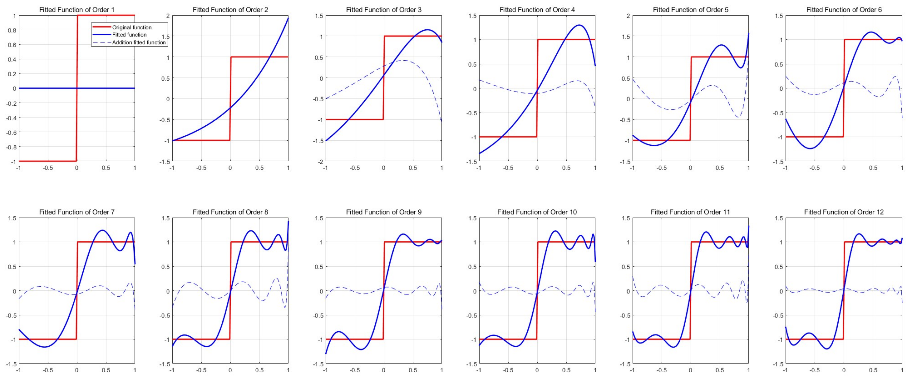

Figure 5: 指数函数拟合示意图

下图展示了利用$\{\ln(x), \ln(x+1), \ln(x+2), \ln(x+3), \dots, \ln(x+n)\}$对数函数基在区间 $[2\pi, 6\pi]$ 拟合 $\sin(x)$ 的示意图，

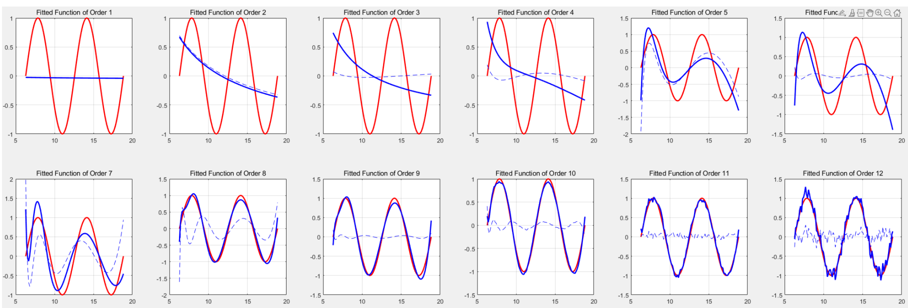

Figure 6: 对数基拟合示意图

此外，我们可以组合使用不同的基，例如利用

$$\{1, x, x^2, x^3, x^4, x^5\} + \{e^x, e^{2x}, e^{3x}, e^{4x}, e^{5x}, e^{6x}\}$$

在区间 $[-\pi, \pi]$ 内拟合 $\sin(x)$。左图显示先使用多项式基进行拟合，在第七项引入指数基；右图则相反。可以看出，若先采用多项式基，其拟合效果更为高效。

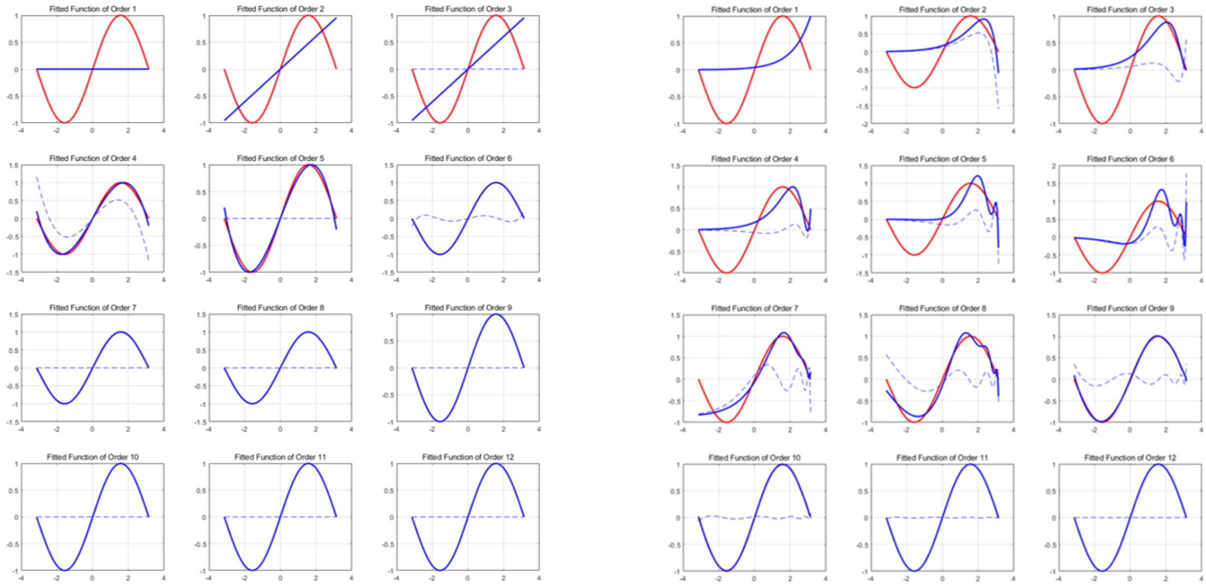

Figure 7: 不同基组合顺序对 $\sin(x)$ 拟合效果的比较

在实际拟合过程中，我们可以任意选择基底 $\{\varphi_n\}_{n\in\mathbf{N}}$ 来张成函数空间 $\mathbf{V}$，再通过投影的方法获得 f 在该基下的投影表示。我们常用基 $\{x^n\}_{n\in\mathbf{N}}$、$\{\cos(n-1)x\}_{n\in\mathbf{N}}$ 、$\{\sin nx\}_{n\in\mathbf{N}}$以及 $\{\cos(n-1)x, \sin nx\}_{n\in\mathbf{N}}$等简单基进行拟合。

3 积分变换

积分变换下的傅里叶形式

函数 f 在空间 V 下的投影为

$$P_V f = \sum_{n=1}^{N} \langle \varphi_n, f \rangle \, \varphi_n.$$

若认为 $P_V f$ 能很好地逼近 f（在紧致区间内一致收敛，在开区间内逐点收敛），则可写成

$$f = \sum_{n=1}^{N} \langle \varphi_n, f \rangle \, \varphi_n.$$

延用上述内积定义，采用正弦基

$$\left\{ \sin \frac{n \cdot 2\pi}{T} x \right\}_{n\in\mathbf{N}},$$

对定义在 $[-\frac{T}{2}, \frac{T}{2}]$ 内的函数 f 进行拟合。首先对该基进行正交归一化，得到

$$\left\{ \sqrt{\frac{2}{T}} \sin \frac{n \cdot 2\pi}{T} x \right\}_{n\in\mathbf{N}}.$$

此时 f 在该空间下的投影为

$$\begin{aligned} P_V f &= \sum_{n=1}^{N} \langle \varphi_n, f \rangle \, \varphi_n \\ &= \frac{2}{T} \sum_{n=1}^{N} \left(\int_{-\frac{T}{2}}^{\frac{T}{2}} f(\tau) \sin \frac{n \cdot 2\pi}{T} \tau \, d\tau \right) \sin \frac{n \cdot 2\pi}{T} x. \end{aligned}$$

利用 $T=\frac{2\pi}{\omega_0}$ 可重写为

$$P_V f=\frac{\omega_0}{\pi}\sum_{n=1}^{N}\left(\int_{-\frac{T}{2}}^{\frac{T}{2}} f(\tau)\sin(n\omega_0\tau)\,d\tau\right)\sin(n\omega_0 t).$$

这便得到了 f 的正弦傅里叶展开。当 $T\to\infty$（即 $\omega_0\to 0$）时，令 $\omega_0=d\omega$ 且 $n\omega_0=\omega$，则

$$P_V f=\frac{1}{\pi}\int_{0}^{N\cdot\omega_0}\left(\int_{-\infty}^{\infty} f(\tau)\sin(\omega\tau)\,d\tau\right)\sin(\omega t)\,d\omega.$$

考虑到外层积分原本从 $\omega_0=0$ 到 $N\omega_0$，为使拟合足够精确，可用无限项进行逼近，上式上限改为 ∞。由于内外层函数均为关于 ω 的奇函数，其乘积为偶函数，因此积分形式亦可调整为

$$P_V f=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\sin(\omega\tau)\,d\tau\right)\sin(\omega t)\,d\omega.$$

这便是傅里叶变换正弦表达式的由来。回顾最初版本

$$P_V f=\sum_{n=1}^{N}\langle\sqrt{\frac{2}{T}}\sin\frac{n\cdot 2\pi}{T}x,f\rangle\sqrt{\frac{2}{T}}\sin\frac{n\cdot 2\pi}{T}x,$$

可将其视为一组奇函数的叠加，因此 f 在 V 上的投影依然为奇函数。

同理，我们可求得 f 在空间

$$W=\mathbf{span}\left\{\cos\frac{n\cdot 2\pi}{T}x\right\}_{0\le n\le N}$$

下的投影：

$$P_W f=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\cos(\omega\tau)\,d\tau\right)\cos(\omega t)\,d\omega.$$

由于正弦基反映 f 的奇函数特性，而余弦基反映其偶函数特性，因此 $P_V f$ 与 $P_W f$ 互相正交。将两者相加，可更全面地表达 f：

$$\begin{aligned}P_V f+P_W f&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\Big[\sin(\omega\tau)\sin(\omega t)+\cos(\omega\tau)\cos(\omega t)\Big]\,d\tau\right)d\omega\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\cos(\omega t-\omega\tau)\,d\tau\right)d\omega.\end{aligned}$$

如果规定 f 为实函数，且 $j^2=-1$，则虚部为0：

$$\frac{j}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\sin(\omega t-\omega\tau)\,d\tau\right)d\omega=0.$$

因此，

$$\begin{aligned}P_V f+P_W f&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\cos(\omega t-\omega\tau)\,d\tau\right)d\omega+\frac{j}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\sin(\omega t-\omega\tau)\,d\tau\right)d\omega\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\Big[\cos(\omega t-\omega\tau)+j\sin(\omega t-\omega\tau)\Big]d\tau\right)d\omega\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)e^{j\omega(t-\tau)}d\tau\right)d\omega\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)e^{-j\omega\tau}d\tau\right)e^{j\omega t}d\omega.\end{aligned}$$

这便得到了标准傅里叶积分表达式：

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)e^{-j\omega\tau}\,d\tau\right)e^{j\omega t}\,d\omega.$$

讨论收敛性

在上述例子中，我们所采用的基均为连续函数，因此 $P_V f + P_W f$ 也为连续函数，可反映 f 的连续性质。在适当条件下，可以证明该投影逐点或一致收敛于 f。

需要注意的是，傅里叶变换也存在局限性。比如将函数投影到正弦函数空间仅能得到其奇函数部分，无法完整反映函数的全部信息。例如，将 $g(x) = x + x^2$ 投影到正弦空间时，通过与各正交归一基求内积得到的系数仅包含 x 项，而 x^2 部分被舍去，从而丢失了相应的信息。

比较系数

在有限维基下的投影过程中，可能出现不同函数具有相同投影的情况，难以区分原始函数。为此，可以采用无限维基，通过比较投影系数来区分不同函数。如果一个函数在无限维空间下的所有投影系数均已确定，则该投影能够唯一标识原始函数。拉普拉斯变换正是基于这一思想来求解微分方程。

积分变换下的拉普拉斯形式

在有限维空间中，在标准正交基 $\{v_i\}_{0 \le i \le n}$ 下的投影，其结果仅为这些基的线性组合，且由于基的线性独立性与唯一性，无论正交化顺序如何，最终得到的投影向量中原始基前面的系数均是相同的。

采用基 $\{e^{-st}\}_{s \in \mathbf{N}}$ 张成函数空间时，我们可以认为相邻的两个s之间的距离为无穷小，这组基张成的是一个足够大的有限维空间。我们可以任意选取其中的 e^{-st} 作为最后一步正交化的对象。除最后一步得到的标准基外，其他标准基均不含 e^{-st} 的线性项，故 e^{-st} 仅在计算最后一步投影中出现。

设最后进行正交化得到的标准基为

$$u_s = c\,e^{-st} + \sum a_{si} u_i,$$

其中 c、a_{si} 与 u_i 均为系统固有性质，与输入的函数 f 无关。函数 f 在 u_s 上获得的投影向量为

$$\langle f, c\,e^{-st} + \sum a_{si} u_i \rangle \Big(c\,e^{-st} + \sum a_{si} u_i \Big),$$

f在这个函数空间下的投影为

$$P_V f = \langle f, c\,e^{-st} + \sum a_{si} u_i \rangle e^{-st} + \sum_{initial}^{final-1} \gamma_{s'} e^{-s't}$$

其中 $\gamma_{s'}$ 与 $\langle f, c\,e^{-st} + \sum a_{si} u_i \rangle$ 分别代表各个 e^{-st} 前面的系数。首项 e^{-st} 的系数中与 f 和 s 均有关的内容为 $\langle f, c\,e^{-st} \rangle$ ，并且其中 c 属于系统固有的属性，与 f 和 s 无关，所以我们可以使用 $\langle f,\, e^{-st} \rangle$ 来反映或者说代表 e^{-st} 的特征投影系数。 在这里定义内积 $\langle f,\, g \rangle = \int_0^\infty f(t) g(t)\, dt$ 那么在这里获得的特征投影系数就为

$$\langle f,\, e^{-st} \rangle = \int_0^\infty f(t) e^{-st}\, dt$$

所以我们可以将拉普拉斯变换理解为一个获得与 e^{-st} 相关系数的过程。同时通过这种过程我们可以获得关系

$$\int_0^\infty f'(t) e^{-st}\, dt = -f(0) + s \int_0^\infty f(t) e^{-st}\, dt$$

建立 f 和 $f^{(n)}$ 之间的关系，进而建立起微分方程 $\sum c_n f^{(n)}(t) = G(t)$ 两边的联系，得到 f 在 e^{-st} 下的特征投影系数，最后通过查表获得对应的原函数。注意同一个特征投影系数有时对应不同的原函数，例如 $\lim_{n\to\infty} \frac{n}{\sqrt{\pi}} e^{-n^2x^2}$ 和 $\lim_{n\to\infty} \frac{n}{\pi} \operatorname{sinc}^2(nx)$ 等 δ 函数的特征投影系数均为 1 。这些 δ 函数的共同特性为 $\int_{-\infty}^{\infty} \delta(x - x_0) f(x)\, dx = f(x_0)$ 。特征投影系数为 1 正是反映了这些函数所具有的这一特性。

推广
上述过程并未利用 $\{e^{-st}\}$ 的特殊函数性质，因此无论是傅里叶变换还是拉普拉斯变换，都适用类似的方法。我们还可以采用如

$$\int_0^\infty f(x)x^s\,dx,$$

（涉及幂函数）或

$$\int_0^\infty f(x)s^x\,dx,$$

（涉及指数函数）或

$$\int_a^b f(x)\ln(x+s)\,dx,$$

（涉及对数函数）等变换，建立与幂、指数、对数等函数相关的关系，从而可以进一步得到 f 与其 n 阶导数 $f^{(n)}$ 之间的联系。

联系前面获得投影的过程:

$$P_V f = \sum_{i=1}^n \langle f, u_i\rangle\, u_i.$$

和傅里叶积分变换的表达式

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^\infty \left(\int_{-\infty}^\infty f(\tau)e^{-j\omega\tau}\,d\tau\right) e^{j\omega t}\,d\omega.$$

我们可以认为做内层积分就是做内积获得投影特征系数的过程，外层积分，就是将对应的系数与正交标准基相乘之后相加的过程。从而将积分变换与投影过程更紧密地链接起来。

March 28, 2025

Abstract

This article systematically explores the theoretical framework and methodological applications of data fitting and integral transforms. The study centers on orthogonal projection in function spaces and explains the role of the Gram–Schmidt orthogonalization in constructing basis functions. Through comparative experiments with typical bases such as polynomials, Fourier bases, and exponential functions, the influence of the choice of basis functions on fitting performance is revealed. Furthermore, the work derives the essence of Fourier transform as an orthogonal projection, extends it to the integral expression of the Laplace transform, and establishes a new perspective for solving differential equations via various integral transforms.

Keywords: Data Fitting; Gram–Schmidt Orthogonalization; Fourier Transform; Laplace Transform; Integral Transforms; Projection in Function Spaces

1 Fitting

Basic Definition:

Fitting refers to the process of finding a function $f(x)$ that approximately describes a given dataset $\{x_i, y_i\}_{i\in\mathbf{N}}$ such that each $f(x_i)$ is as close as possible to the corresponding y_i. A common measure for the fitting performance is the sum of squared residuals

$$\sum_{i=1}^{N} |f(x_i) - y_i|^2.$$

Similarly, to compare the difference between two functions $f(x)$ and $g(x)$, a similar criterion can be used, whose continuous form is written as

$$\int_a^b |f(x) - g(x)|^2 \, dx.$$

In the vector space of functions, we define the inner product as

$$\langle f, g\rangle = \int_a^b f(x)g(x)\, dx,$$

and then define the norm

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

so that $\|f - g\|$ naturally represents the distance between f and g.

How to Perform Fitting

In high school, we learned to fit data points using the least squares method, but this method is only applicable to linear relationships, which are relatively rare in nature.

The least squares method uses an equation of the form

$$y = \hat{a} \cdot 1 + \hat{b} \cdot x,$$

taking $\{1, x\}$ as the basis to fit the data. Later, through Taylor expansion, we began using $\{1, x, x^2, x^3, \dots\}$ as basis functions to approximate functions. Furthermore, we introduced the Fourier series

$$\{1, \cos x, \cos 2x, \cos 3x, \dots, \sin x, \sin 2x, \sin 3x, \dots\}$$

as a fitting tool.

For example, consider $f(x) = 3x^2 + 2x - 1$. If the least squares method is used, it can only capture the general trend over a finite interval; however, if the Taylor series is used, the first three terms can fully represent the information contained in $f(x)$, yielding precise results in both inversion and prediction. Conversely, if the Fourier series is applied, then within a given closed interval

$$\sum_{i=-n}^{n} C_i e^{j\omega_i x}$$

converges absolutely to $f(x) = 3x^2 + 2x - 1$; but outside that interval, since the Fourier series is composed of sine and cosine functions exhibiting clear periodicity while $f(x)$ is non-periodic, the fitting fails. This illustrates the differences arising from the choice of fitting bases. Although we cannot always select the perfect basis for prediction, using an appropriate basis within a given range can effectively reflect the variation in the data or function.

2 Gram–Schmidt Orthogonalization

Method for Function Fitting

We use a set of basis functions $\{v_1(x), v_2(x), v_3(x), \dots, v_n(x)\}$ to fit a target function. Any function in the function space $\mathbf{V}$ spanned by this set can be represented as

$$u(x) = \sum_{i=1}^{n} C_i v_i(x).$$

Our goal is to determine a $u(x)$ such that the sum of squared residuals

$$\sum_{i=1}^{N} |f(x_i) - u(x_i)|^2$$

is minimized. A weight function $r(x)$ can also be introduced to account for the different importance of data points, in which case the sum of squared residuals becomes

$$\sum_{i=1}^{N} |f(x_i) - u(x_i)|^2 r(x)$$

or its continuous form

$$\int_a^b |f(x) - u(x)|^2 r(x)\, dx.$$

For simplicity, we usually take $r(x) = 1$. Thus, the expression

$$\int_a^b |f(x) - u(x)|^2\, dx$$

reflects the distance between f and u in the overall space $\mathbf{W}$. To minimize

$$\int_a^b |f(x) - u(x)|^2\, dx,$$

u must be the orthogonal projection of f onto the space $\mathbf{V}$.

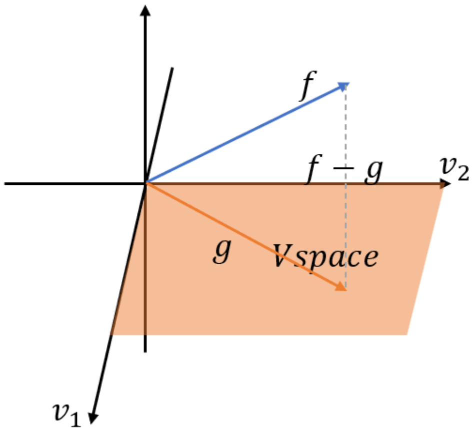

Figure 8: Schematic of Orthogonal Projection

Constructing an Orthonormal Basis

In a vector space equipped with an inner product, we can use the Gram–Schmidt orthogonalization process to convert the basis $\{v_1(x), v_2(x), v_3(x), \ldots, v_n(x)\}$ into an orthonormal basis $\{u_1(x), u_2(x), u_3(x), \ldots, u_n(x)\}$:

$$\textbf{Step 1:}\quad u_1(x) = \frac{v_1(x)}{\sqrt{\langle v_1, v_1\rangle}} = \frac{v_1(x)}{\sqrt{\int_a^b v_1(x)^2\,dx}},$$

$$\textbf{Step 2:}\quad u_2'(x) = v_2(x) - \langle v_2, u_1\rangle\, u_1(x), \quad u_2(x) = \frac{u_2'(x)}{\sqrt{\langle u_2', u_2'\rangle}},$$

$$\vdots$$

$$\textbf{Step n:}\quad u_n'(x) = v_n(x) - \sum_{i=1}^{n-1}\langle v_n, u_i\rangle\, u_i(x), \quad u_n(x) = \frac{u_n'(x)}{\sqrt{\langle u_n', u_n'\rangle}}.$$

It is easy to verify that $\langle u_i, u_j\rangle = \delta_{ij}$. For example:

$$\textbf{Example:}\quad \langle u, v\rangle = \int_{-1}^{1} u(x)v(x)\,dx,$$

$$\{1, x, x^2\} \Rightarrow \left\{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}\,x}{\sqrt{2}}, \frac{3\sqrt{10}}{4}x^2 - \frac{\sqrt{10}}{4}\right\},$$

$$\{1, \cos\pi x, \cos 2\pi x\} \Rightarrow \left\{\frac{1}{\sqrt{2}}, \cos\pi x, \cos 2\pi x\right\}.$$

Obtaining the Projection

With the orthonormal basis at hand, we can project f onto $\mathbf{V}$, with the projection denoted by $P_V f$:

$$P_V f = \sum_{i=1}^{n}\langle f, u_i\rangle\, u_i.$$

In this way, the distance between f and the space $\mathbf{V}$ (i.e. the sum of squared residuals $\int_a^b |f(x) - P_V f(x)|^2\,dx$) is minimized.

For example, when we use the basis $\{\sin \pi x, \sin 2\pi x, \sin 3\pi x, \ldots \sin n\pi x\}$ to fit the function

$$u(x) = \begin{cases} 0, & -1 < x < 0, \\ 1, & 0 \le x < 1, \end{cases}$$

it is clear that this is exactly the process of a Fourier series expansion; because the terms are orthogonal, each newly added term exhibits the shape of a sine function.

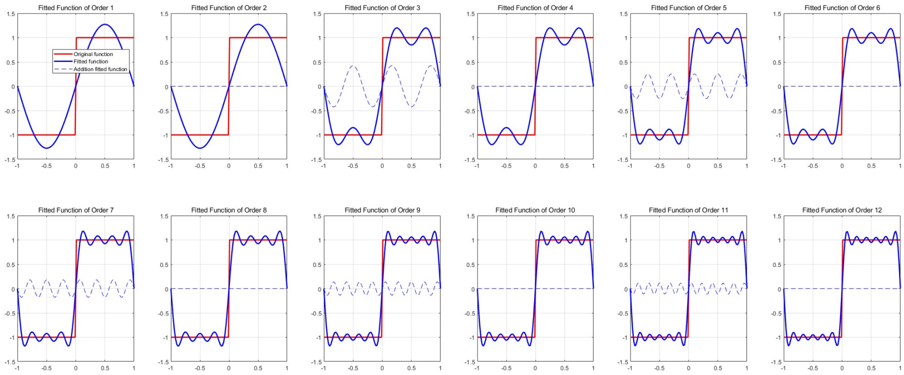

Figure 9: Schematic of Sine Series Fitting

The figure below shows the expansion effect when using the basis

$$\left\{ \sin\!\Big(\Big(\frac{1 - (-1)^n + 2n}{4} \Big) \pi x + \frac{\pi\,(1 + (-1)^n)}{4} \Big) \right\}_{0 \le n \le N}$$

(i.e. a basis composed of alternating cosine and sine functions). It can be seen that for the cosine terms, the additional components are all zero because the cosine function is even while the square wave to be fitted is odd; on a symmetric interval they are naturally orthogonal, resulting in a zero projection. This shows that the cosine functions cannot reflect the characteristics of the square wave, and we need to supplement with a set of odd function bases to fit the target function.

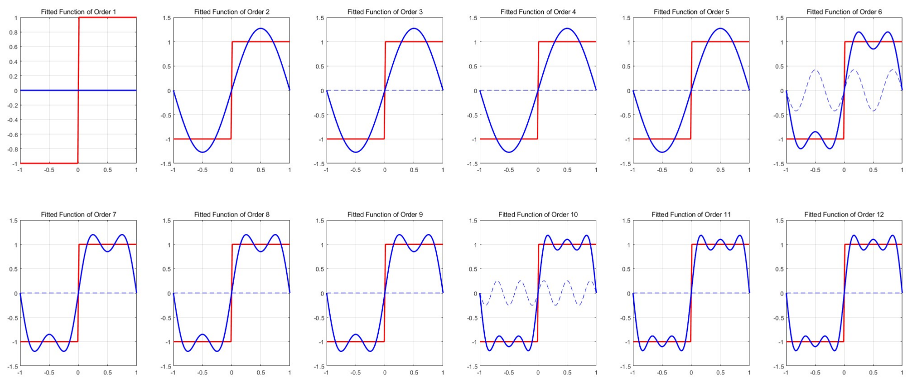

Figure 10: Fourier Expansion with Alternating Cosine and Sine Bases

Similarly, when using the basis $\{1, x, x^2, x^3, \ldots x^n\}$ (after orthonormalization) for fitting, the first few terms exhibit obvious polynomial characteristics, while the later terms, being linear combinations of several polynomial functions, display ambiguous polynomial characteristics.

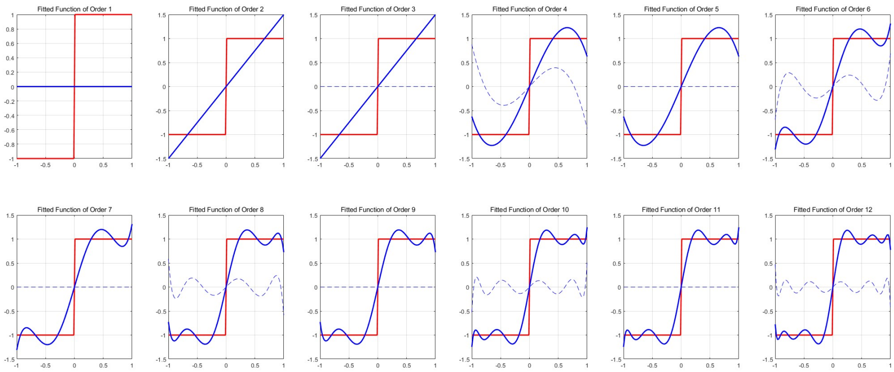

Figure 11: Schematic of Polynomial Basis Fitting

In addition, other functions can be used for fitting, for example the basis $\{1, e^x, e^{2x}, e^{3x}, \ldots e^{nx}\}$ (after orthonormalization).

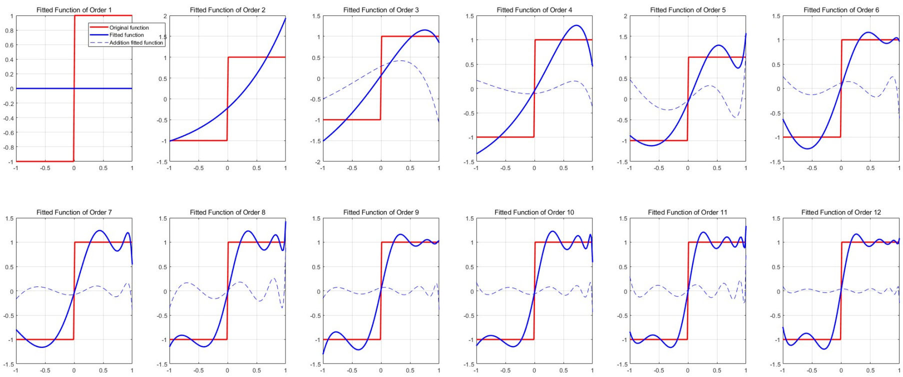

Figure 12: Schematic of Exponential Function Fitting

The figure below shows the fitting of $\sin(x)$ on the interval $[2\pi, 6\pi]$ using a logarithmic basis

$$\{\ln(x), \ln(x+1), \ln(x+2), \ln(x+3), \ldots, \ln(x+n)\}.$$

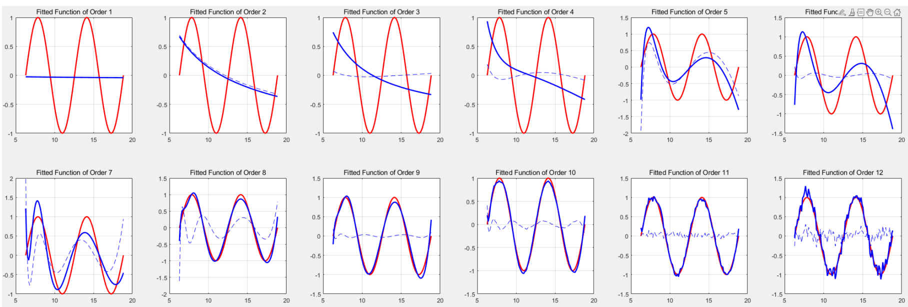

Figure 13: Schematic of Logarithmic Basis Fitting

Moreover, we can combine different bases. For example, using

$$\{1, x, x^2, x^3, x^4, x^5\} + \{e^x, e^{2x}, e^{3x}, e^{4x}, e^{5x}, e^{6x}\}$$

to fit $\sin(x)$ over the interval $[-\pi, \pi]$. The left figure shows the fitting process when the polynomial basis is used first and the exponential basis is introduced at the seventh term; the right figure shows the reverse order. It can be seen that if the polynomial basis is used first, the fitting effect is more efficient.

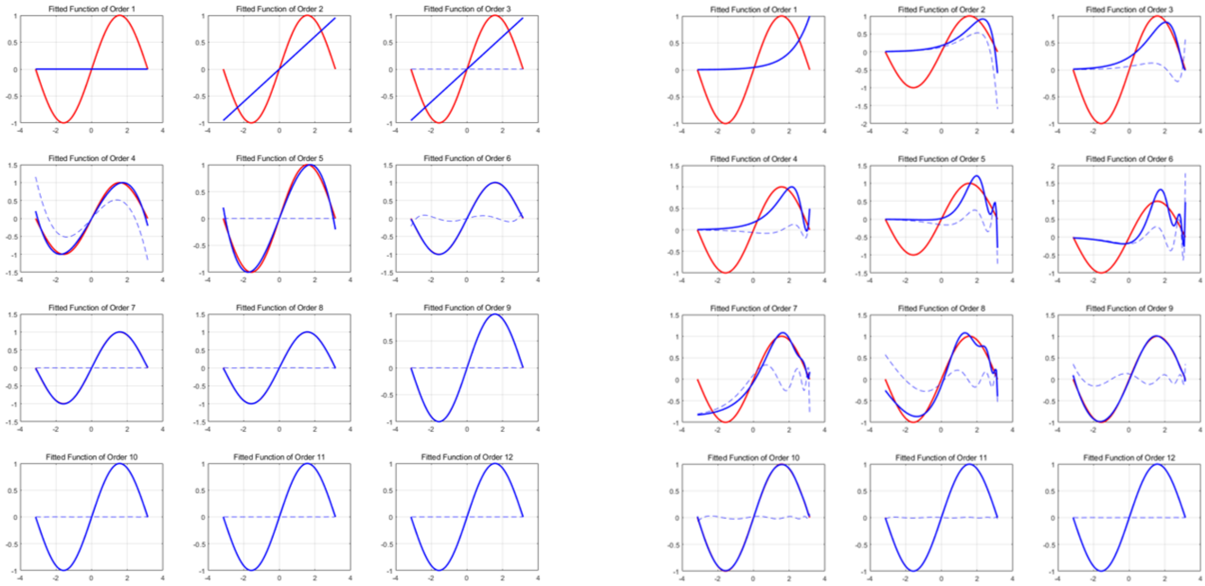

Figure 14: Comparison of $\sin(x)$ Fitting with Different Basis Combination Orders

In practical fitting, we can choose any basis $\{\varphi_n\}_{n\in\mathbf{N}}$ to span the function space $\mathbf{V}$, and then obtain the projection representation of f on this basis via the projection method. Commonly used bases include $\{x^n\}_{n\in\mathbf{N}}$, $\{\cos(n-1)x\}_{n\in\mathbf{N}}$, $\{\sin nx\}_{n\in\mathbf{N}}$, and $\{\cos(n-1)x, \sin nx\}_{n\in\mathbf{N}}$.

3 Integral Transforms

Fourier Form under Integral Transforms

The projection of the function f onto the space V is given by

$$P_V f = \sum_{n=1}^{N} \langle \varphi_n, f\rangle \, \varphi_n.$$

If we assume that $P_V f$ approximates f well (uniformly convergent on a compact interval and pointwise convergent on an open interval), then it can be written as

$$f = \sum_{n=1}^{N} \langle \varphi_n, f\rangle \, \varphi_n.$$

Using the aforementioned inner product definition, we fit the function f defined on $[-\frac{T}{2}, \frac{T}{2}]$ using the sine basis

$$\left\{ \sin \frac{n \cdot 2\pi}{T} x \right\}_{n\in\mathbf{N}},$$

and first orthonormalize this basis to obtain

$$\left\{ \sqrt{\frac{2}{T}} \sin \frac{n \cdot 2\pi}{T} x \right\}_{n\in\mathbf{N}}.$$

Then the projection of f in this space is

$$\begin{aligned} P_V f &= \sum_{n=1}^{N} \langle \varphi_n, f \rangle \, \varphi_n \\ &= \frac{2}{T} \sum_{n=1}^{N} \left(\int_{-\frac{T}{2}}^{\frac{T}{2}} f(\tau) \sin \frac{n \cdot 2\pi}{T} \tau \, d\tau \right) \sin \frac{n \cdot 2\pi}{T} x. \end{aligned}$$

Using $T = \frac{2\pi}{\omega_0}$, this can be rewritten as

$$P_V f = \frac{\omega_0}{\pi} \sum_{n=1}^{N} \left(\int_{-\frac{T}{2}}^{\frac{T}{2}} f(\tau) \sin(n\omega_0 \tau) \, d\tau \right) \sin(n\omega_0 t).$$

This is the sine Fourier expansion of f. When $T \to \infty$ (i.e. $\omega_0 \to 0$), letting $\omega_0 = d\omega$ and $n\omega_0 = \omega$, we have

$$P_V f = \frac{1}{\pi} \int_0^{N \cdot \omega_0} \left(\int_{-\infty}^{\infty} f(\tau) \sin(\omega \tau) \, d\tau \right) \sin(\omega t) \, d\omega.$$

Since the outer integral originally runs from $\omega_0 = 0$ to $N\omega_0$, to achieve sufficient fitting accuracy, we approximate with an infinite number of terms and change the upper limit to ∞. Because the inner and outer functions are odd functions of ω and their product is even, the integral form can also be adjusted to

$$P_V f = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\tau) \sin(\omega \tau) \, d\tau \right) \sin(\omega t) \, d\omega.$$

This is the origin of the sine expression of the Fourier transform. Recalling the initial version

$$P_V f = \sum_{n=1}^{N} \langle \sqrt{\frac{2}{T}} \sin \frac{n \cdot 2\pi}{T} x, f \rangle \sqrt{\frac{2}{T}} \sin \frac{n \cdot 2\pi}{T} x,$$

we can view it as a superposition of odd functions; therefore, the projection of f onto V is still an odd function.

Similarly, we can obtain the projection of f onto the space

$$W = \textbf{span} \left\{ \cos \frac{n \cdot 2\pi}{T} x \right\}_{0 \le n \le N},$$

which is

$$P_W f = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\tau) \cos(\omega \tau) \, d\tau \right) \cos(\omega t) \, d\omega.$$

Since the sine basis reflects the odd part of f and the cosine basis reflects its even part, $P_V f$ and $P_W f$ are orthogonal to each other. By adding the two, we obtain a more complete representation of f:

$$\begin{aligned} P_V f + P_W f &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\tau) \Big[\sin(\omega\tau) \sin(\omega t) + \cos(\omega\tau) \cos(\omega t) \Big] \, d\tau \right) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\tau) \cos(\omega t - \omega \tau) \, d\tau \right) d\omega. \end{aligned}$$

If we assume f is a real function and $j^2 = -1$, then the imaginary part is zero:

$$\frac{j}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\tau) \sin(\omega t - \omega \tau) \, d\tau \right) d\omega = 0.$$

Therefore,

$$
\begin{aligned}
P_V f + P_W f &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\cos(\omega t-\omega\tau)\,d\tau\right)d\omega + \frac{j}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\sin(\omega t-\omega\tau)\,d\tau\right)d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)\Big[\cos(\omega t-\omega\tau)+j\sin(\omega t-\omega\tau)\Big]d\tau\right)d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)e^{j\omega(t-\tau)}d\tau\right)d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)e^{-j\omega\tau}d\tau\right)e^{j\omega t}d\omega.
\end{aligned}
$$

This yields the standard Fourier integral expression:

$$
f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(\tau)e^{-j\omega\tau}\,d\tau\right)e^{j\omega t}\,d\omega.
$$

Discussion of Convergence

In the above examples, the bases we employed are all continuous functions, so $P_V f + P_W f$ is also continuous, reflecting the continuity of f. Under appropriate conditions, it can be proven that the projection converges to f either pointwise or uniformly.

It is important to note that the Fourier transform also has limitations. For instance, projecting a function onto the sine function space only retrieves its odd part and cannot fully capture all the information of the function. For example, when projecting $g(x) = x + x^2$ onto the sine space, the coefficients obtained by taking the inner products with the orthonormal bases only contain the x term, while the x^2 component is discarded, leading to loss of the corresponding information.

Comparison of Coefficients

In the projection process using a finite-dimensional basis, different functions may have the same projection, making it difficult to distinguish the original functions. To resolve this, an infinite-dimensional basis can be used to distinguish different functions by comparing their projection coefficients. If all the projection coefficients of a function in an infinite-dimensional space are determined, then the projection uniquely identifies the original function. The Laplace transform is based on this idea for solving differential equations.

Laplace Form under Integral Transforms

In a finite-dimensional space, the projection under a standard orthonormal basis $\{v_i\}_{0\le i\le n}$ results in a linear combination of these bases, and due to the linear independence and uniqueness of the bases, regardless of the order of orthogonalization, the coefficients for the original bases in the final projection vector remain the same.

When using the basis $\{e^{-st}\}_{s\in\mathbf{N}}$ to span a function space, we can think of the distance between adjacent s values as infinitesimal, so that the basis spans a sufficiently large finite-dimensional space. We can arbitrarily select one of the e^{-st} terms as the object of the final orthogonalization. Except for the final standard basis obtained in the last step, the other standard bases do not contain the linear term of e^{-st}; hence, e^{-st} appears only in the final projection computation.

Let the standard basis obtained from the final orthogonalization be

$$
u_s = c\,e^{-st} + \sum a_{si}u_i,
$$

where c, a_{si}, and u_i are inherent properties of the system and are independent of the input function f. The projection of f onto u_s is

$$\langle f, c\,e^{-st} + \sum a_{si} u_i \rangle \Big(c\,e^{-st} + \sum a_{si} u_i \Big),$$

and the projection of f in this function space is

$$P_V f = \langle f, c\,e^{-st} + \sum a_{si} u_i \rangle e^{-st} + \sum_{initial}^{final-1} \gamma_{s'} e^{-s't},$$

where $\gamma_{s'}$ and $\langle f, c\,e^{-st} + \sum a_{si} u_i \rangle$ represent the coefficients in front of each e^{-st}. In the coefficient of the first term e^{-st}, the part related to both f and s is $\langle f, c\,e^{-st} \rangle$, and since c is an inherent attribute of the system independent of f and s, we can use $\langle f,\, e^{-st} \rangle$ to reflect or represent the characteristic projection coefficient of e^{-st}. Here we define the inner product $\langle f,\, g \rangle = \int_0^\infty f(t) g(t)\, dt$. Thus, the characteristic projection coefficient obtained is

$$\langle f,\, e^{-st} \rangle = \int_0^\infty f(t) e^{-st}\, dt.$$

So, we can interpret the Laplace transform as a process that obtains the coefficient related to e^{-st}. Through this process, we obtain the relation

$$\int_0^\infty f'(t) e^{-st}\, dt = -f(0) + s \int_0^\infty f(t) e^{-st}\, dt,$$

which establishes the relationship between f and its derivatives $f^{(n)}$, and thus connects the two sides of a differential equation such as $\sum c_n f^{(n)}(t) = G(t)$. This gives the characteristic projection coefficients of f under e^{-st}, and finally, by consulting tables, the corresponding original function is obtained. Note that sometimes the same characteristic projection coefficient corresponds to different original functions; for example, the delta function projection coefficients of $\lim_{n\to\infty} \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}$ and $\lim_{n\to\infty} \frac{n}{\pi} \operatorname{sinc}^2(nx)$ are both 1. The common property of these delta functions is that $\int_{-\infty}^{\infty} \delta(x - x_0) f(x)\, dx = f(x_0)$. A projection coefficient of 1 reflects precisely this property of the functions.

Generalization

The above process does not utilize the special functional properties of $\{e^{-st}\}$, so similar methods apply to both the Fourier transform and the Laplace transform. We can also employ transforms such as

$$\int_0^\infty f(x) x^s\, dx,$$

(involving power functions), or

$$\int_0^\infty f(x) s^x\, dx,$$

(involving exponential functions), or

$$\int_a^b f(x) \ln(x + s)\, dx,$$

(involving logarithmic functions) to establish relations related to power, exponential, or logarithmic functions, thereby further obtaining the relationship between f and its nth derivative $f^{(n)}$.

Recall the projection process we employed:

$$P_V f = \sum_{i=1}^{n} \langle f, u_i \rangle \, u_i.$$

and the expression for the Fourier integral transform:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(\tau) e^{-j\omega\tau} \, d\tau \right) e^{j\omega t} \, d\omega.$$

We can consider the inner integral as the process of taking an inner product to obtain the projection coefficients, and the outer integral as the process of summing the products of these coefficients with the orthonormal basis. In this way, the integral transform is closely linked with the projection process.

1 Code Support

```
n = 12;              % Fitting order
a = 2*pi; b = 6*pi;  % Interval
orifun = @(x) sin(x);% Function to be fitted
createfun = @(ii,x) log(x+ii-1);
% Options for basis:
% Exponential: exp((ii-1)*x)
% Polynomial: x.^(ii-1)
% Sine: sin(ii*2*pi/(b-a)*(x-(a+b)/2))
% Cosine: cos((ii-1)*2*pi/(b-a)*(x-(a+b)/2))
% Logarithm: log(ii+x)

x = linspace(a,b,100);

fun = cell(1,n);       % Basis functions (not orthonormalized)
normalfun = cell(1,n); % Orthonormalized basis functions

% Generate the required bases
for ii = 1:n
    fun{ii} = @(x) createfun(ii,x);
    normalfun{ii} = fun{ii};
end

% Normalize the first vector
coeff = integral(@(x) normalfun{1}(x).^2, a, b);
normalfun{1} = @(x) normalfun{1}(x) / sqrt(coeff);
% Orthonormalize the n vectors
for ii = 2:n
    for jj = 1:ii-1
        coeff = integral(@(x) normalfun{ii}(x) .* normalfun{jj}(x), a, b
           );
        normalfun{ii} = @(x) normalfun{ii}(x) - coeff .* normalfun{jj}(x
           );
    end
    coeff = integral(@(x) normalfun{ii}(x).^2, a, b);
    normalfun{ii} = @(x) normalfun{ii}(x) / sqrt(coeff);
end

% Fitting process
coeff = zeros(1, n);
fitfun = @(x) 0;

% Create figure layout
figure;
rows = ceil(n/6);
cols = 6;
% Iterative fitting
for ii = 1:n
    coeff(ii) = integral(@(x) orifun(x) .* normalfun{ii}(x), a, b);
    fitfun = @(x) fitfun(x) + coeff(ii) .* normalfun{ii}(x);
    % Plot each fitting curve
    subplot(rows, cols, ii);
    plot(x, orifun(x), 'r', 'LineWidth', 2, 'DisplayName', 'Original␣
       function');
    hold on;
    plot(x, fitfun(x), 'b', 'LineWidth', 2, ...
         'DisplayName', ['Fitted␣function␣(Order␣' num2str(ii) ')']);
```

```
    contribution = @(x) coeff(ii) .* normalfun{ii}(x);
    plot(x, contribution(x), '--b', 'DisplayName', ...
         ['Additional fitted function (Order ' num2str(ii) ')']);
    % Set title and legend for each subplot
    title(['Fitted Function of Order ' num2str(ii)]);
    grid on;
end
```