

第一章 概论

Home Page

Title Page



Page 2 of 66

Go Back

Full Screen

Close

Quit

从 Fourier 分析到小波分析



Jean Baptiste Joseph Fourier (1768~1830),
法国数学家、物理学家. 由于对热传导理论的
贡献, 于1817年当选为法国科学院院士.
其经典著作《热的解析理论》对19世纪数
学和理论物理学的发展产生了深远的影响.

[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 3 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

- Fourier 之后，经过 Dirichlet、Riemann、Lipschitz 以及 Jordan 等众多数学家的努力，不仅弥补了 Fourier 工作中的不足，而且极大的发展了以 Fourier 命名的级数理论，扩大了 Fourier 分析的应用范围，还使得这一理论成为研究周期现象不可缺少的工具。
- Fourier 分析在概念和方法上对其他数学分支的发展给与了深刻的影响。现代正确的函数概念是由 Dirichlet 在研究三角级数收敛性的论文中首先提出的；Riemann 积分的定义是由 Riemann 在题为《用三角级数来表示函数》的论文中明确引进的；Cantor 在对三角级数的唯一性集合的研究中奠定了点集论的基础；Fourier 分析的研究还推动了函数空间理论的发展。

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 4 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

- Fourier 分析最重要的意义是引进了频率的概念，把一个函数展开成各种频率的谐波的线性叠加，由此引出了一系列频谱分析的理论，使对函数性态的研究转换成对 Fourier 系数和 Fourier 变换的研究。
- 长期以来，Fourier 分析不论在数学还是在工程中一直占据着极其重要的地位。然而，对于信号尤其是非平稳信号的分析，Fourier 分析理论存在着缺陷。

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 5 of 66

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

Fourier 分析方法的缺陷

- 三角基函数作为一定周期、一定波形的光滑函数，对于存在间断点的信号进行近似时会出现 Gibbs 现象，对于一般的非周期信号三角基近似经常不是最优选择。
- 分辨率不高，由于三角基函数的频谱点等距分布，因此对于具有突变性质的非平稳信号更是显得无能为力。
- 三角基函数在时域上不能局部化，无法实现时域的局部分析。
- 经典的 Fourier 积分不能同时进行时域和频域分析。

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [▶](#)[Page 6 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

小波分析的思想

寻找小的波或”小波”, 用其整数平移和二进伸缩生成整个 $L^2(\mathbb{R})$ 空间.

定义 一个函数 $\psi \in L^2(\mathbb{R})$ 称为正交小波, 如果其整数平移和二进伸缩

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), \quad j, k \in \mathbb{Z},$$

构成 $L^2(\mathbb{R})$ 的标准正交基. 即 $\langle \psi_{j,k}, \psi_{l,m} \rangle = \delta_{j,l} \delta_{k,m}$, $j, k, l, m \in \mathbb{Z}$,

且任意的 $f \in L^2(\mathbb{R})$ 可写成

$$f(t) = \sum_{j,k=-\infty}^{+\infty} c_{j,k} \psi_{j,k}(t).$$

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 7 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

小波分析的兴起

- 1910年，Haar 提出了 Haar 函数系（Haar基），它是最早出现的小波正交基。
- 1936年，Littlewood 和 Paley 对 Fourier 级数建立了 Littlewood-Paley 理论，给出了频率按二进位划分的思想，为小波的发展奠定了理论基础。
- 20世纪六七十年代，Calderon 和 Coifman 建立了对 H_p 空间的原子分解，其离散形式已接近于小波展开。
- 1984年，法国地球物理学家 Morlet 在通过分析由爆炸方法产生的人造地震数据来探知地下岩石油层分布时首次提出了小波变换的概念。

Home Page

Title Page

◀ ▶

◀ ▶

Page 8 of 66

Go Back

Full Screen

Close

Quit

- 1986年，Meyer 构造了具有一定衰减性的光滑小波函数，并与从事计算机视觉和图像分析的专家 Mallat 合作，建立了多分辨率分析的理论框架，给出了构造小波基的一般化方法。
- 1989年，Mallat 在塔式分解算法的启发下，建立了与 FFT 相应的小波变换的快速算法-Mallat 算法，实现了小波分析从数学到技术的转变。
- 1988年，Daubechies 给出了具有任意阶消失矩和光滑性的紧支集小波的构造方法，按照这种方法构造得到的小波基称为 Daubechies 小波族，该族小波得到了非常广泛的应用。

Home Page

Title Page

◀ ▶

◀ ▶

Page 9 of 66

Go Back

Full Screen

Close

Quit

小波分析的优势

- 小波基具有良好的局部化性质, 便于做局部分析;
- 多分辨功能—数学显微镜. 小波基函数可按分辨率伸缩, 波形可窄可宽. 可以按分辨率聚焦到研究对象的任意细节, 具有很强的分辨功能;
- 由于小波基的平移, 伸展功能, 使小波具有灵活可变的时间—频率窗. 在高频时, 时间窗变窄, 而在低频时, 时间窗变宽. 这有利于分析非平稳信号;
- 小波中的伸缩率与计算机视觉及人眼视觉特征相吻合, 有利于图像数据的压缩.

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 10 of 66

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

- 从数学的角度看，小波分析是在 Fourier 分析的基础上发展起来的数学分支，它是 Fourier 分析、泛函分析、数值分析的完美结晶。
- 从信息科学的角度看，小波分析是时频分析、多尺度分析的进一步发展，已经成为信号分析和信息处理的新的强有力的工具。

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 11 of 66

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

小波的应用研究

- 小波在数学其它分支中的应用，如微分方程、积分方程、函数逼近、分形、混沌、概率小波、非线性分析等.
- 小波在信号处理中的应用，包括信号检测、目标识别及去噪等.
- 小波在图像处理中的应用，包括图像数据压缩、去噪、数字水印、指纹鉴别、模式识别等.
- 小波在通信中的应用，如在CDMA、自适应均衡、扩频通信和分形调制等方面的应用.

Home Page

Title Page

◀ ▶

◀ ▶

Page 12 of 66

Go Back

Full Screen

Close

Quit

内积空间

定义： 设 V 是一个复的线性空间，函数 $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ 称为 V 上的内积，如果满足下列条件：对任意的 $u, v, w \in V$ 及 $c \in \mathbb{C}$ 有

- **正定性:** $\langle v, v \rangle \geq 0$, 且 $\langle v, v \rangle = 0$ 当且仅当 $v = 0$;
- **共轭对称性:** $\langle v, w \rangle = \overline{\langle w, v \rangle}$;
- **齐次性:** $\langle cv, w \rangle = c\langle v, w \rangle$;
- **可加性:** $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.

定义了内积的线性空间称为内积空间.

Home Page

Title Page

◀ ▶

◀ ▶

Page 13 of 66

Go Back

Full Screen

Close

Quit

内积的其他性质

- 双线性性:

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$$

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$$

- 共轭齐次性:

$$\langle v, cw \rangle = \bar{c} \langle v, w \rangle$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 14 of 66

Go Back

Full Screen

Close

Quit

例1 在 n 维复线性空间 \mathbb{C}^n 中, 对任意的 $v = (v_1, v_2, \dots, v_n)$, $w = (w_1, w_2, \dots, w_n) \in \mathbb{C}^n$, 定义内积

$$\langle v, w \rangle = \sum_{j=1}^n v_j \overline{w_j}.$$

则 \mathbb{C}^n 成为一个内积空间。

在 n 维实线性空间 \mathbb{R}^n 中, 对任意的 $v = (v_1, v_2, \dots, v_n)$, $w = (w_1, w_2, \dots, w_n) \in \mathbb{R}^n$, 定义内积

$$\langle v, w \rangle = \sum_{j=1}^n v_j w_j.$$

则 \mathbb{R}^n 成为一个内积空间。

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 15 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

例2 设 $v = (v_1, v_2), w = (w_1, w_2) \in \mathbb{C}^2$. 在 \mathbb{C}^2 中定义内积

$$\langle v, w \rangle = (\overline{w_1}, \overline{w_2}) \begin{pmatrix} 2 & -i \\ i & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

则 \mathbb{C}^2 成为一个内积空间. 更一般地, 定义内积

$$\langle v, w \rangle = (\overline{w_1}, \overline{w_2}) A \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

其中 A 是正定的 Hermitian 矩阵.

- 正定性 $\Rightarrow \langle v, v \rangle = (\overline{v_1}, \overline{v_2}) A \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} > 0$
- Hermitian \Rightarrow

$$\overline{\langle w, v \rangle} = (\overline{w_1}, \overline{w_2}) A^* \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = (\overline{w_1}, \overline{w_2}) A \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \langle v, w \rangle$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 16 of 66

Go Back

Full Screen

Close

Quit

范数与距离

- 在内积空间中定义一个与内积有关的范数: 称之为诱导范数

$$\|v\| := \sqrt{\langle v, v \rangle}$$

- 根据诱导范数, 自然可以引入内积空间中任意两个元素 v, w 之间的距离:

$$d(v, w) := \|v - w\|$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 17 of 66

Go Back

Full Screen

Close

Quit

收敛

- **收敛点列** 设 V 是内积空间, $v_k, k = 1, 2, \dots$ 是 V 中的点列且 $v \in V$. 如果当 $k \rightarrow \infty$ 时, $\|v_k - v\| \rightarrow 0$, 则称点列 $v_k, k = 1, 2, \dots$ 收敛于 v , 记为 $v_k \rightarrow v, k \rightarrow \infty$.
- **Cauchy 点列** 设 V 是内积空间, $v_k, k = 1, 2, \dots$ 是 V 中的点列. 如果 $\forall \epsilon > 0$, 存在正整数 N , 使得当 $m, n > N$ 时, $\|v_m - v_n\| < \epsilon$, 则称 $v_k, k = 1, 2, \dots$ 是 V 中的 Cauchy 点列.
- **Hilbert 空间** 如果内积空间 V 中每一个 Cauchy 点列都是收敛点列, 则称 V 是完备的. 完备的内积空间称为 Hilbert 空间.

Home Page

Title Page

◀ ▶

◀ ▶

Page 18 of 66

Go Back

Full Screen

Close

Quit

L^2 空间

$L^2([a, b])$ 表示 $[a, b]$ 上所有平方可积函数的全体, 即

$$L^2([a, b]) = \{f : [a, b] \rightarrow \mathbb{C}, \int_a^b |f(t)|^2 dt < \infty\}.$$

- $L^2([a, b])$ 是 \mathbb{C} 上的一个线性空间.
- $L^2([a, b])$ 无限维的线性空间. 例如, $\{1, t, t^2, t^3, \dots\}$ 是 $L^2([0, 1])$ 中的线性无关组.
- 在 $L^2([a, b])$ 中定义内积: $\forall f, g \in L^2([a, b])$

$$\langle f, g \rangle_{L^2} = \int_a^b f(t) \overline{g(t)} dt,$$

则 $L^2([a, b])$ 构成一个 Hilbert 空间.

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 19 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

L^2 内积的正定性

- 如果 $0 = \langle f, f \rangle = \int_a^b |f(t)|^2 dt$, 并且 f 是连续函数, 则对所有的点 t 有 $f(t) = 0$.

证明: 假设存在 $t_0 \in [a, b]$ 使得 $|f(t_0)| > 0$. 由 f 在 t_0 的连续性可知, 存在 $\delta > 0$, 使得对任意的 $t \in [t_0 - \delta, t_0 + \delta]$ 有

$$|f(t) - f(t_0)| < \frac{1}{2}|f(t_0)|,$$

即 $|f(t)| > \frac{1}{2}|f(t_0)|, \forall t \in [t_0 - \delta, t_0 + \delta]$. 于是, 我们得到

$$\int_a^b |f(t)|^2 dt \geq \int_{t_0 - \delta}^{t_0 + \delta} |f(t)|^2 dt \geq 2\delta \cdot \frac{1}{4}|f(t_0)|^2 > 0.$$

这与 $\langle f, f \rangle = 0$ 矛盾.

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 20 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

- 如果 $0 = \langle f, f \rangle = \int_a^b |f(t)|^2 dt$, 并且 f 有有限个间断点, 则除有限个点外 $f(t) = 0$.
- 如果 $f \in L^2([a, b])$ 并且 $0 = \langle f, f \rangle = \int_a^b |f(t)|^2 dt$, 则在 $[a, b]$ 上除零测度集外 $f(t) = 0$.
- 在 $[a, b]$ 上除零测度集外 $f(t) = g(t)$, 则认为函数 f 和 g 相等. 在此意义下, L^2 内积的正定性成立.

Home Page

Title Page

◀ ▶

◀ ▶

Page 21 of 66

Go Back

Full Screen

Close

Quit

l^2 空间

l^2 表示所有满足

$$\sum_{n=-\infty}^{\infty} |x_n|^2 < \infty$$

的序列 $X = \cdots, x_{-1}, x_0, x_1, \cdots, x_i \in \mathbb{C}$ 的集合. 对 l^2 中任意序列

$$X = \cdots, x_{-1}, x_0, x_1, \cdots,$$

$$Y = \cdots, y_{-1}, y_0, y_1, \cdots,$$

定义内积

$$\langle X, Y \rangle_{l^2} = \sum_{n=-\infty}^{\infty} x_n \overline{y_n}.$$

则 l^2 空间构成无限维的 Hilbert 空间.

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 22 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

L^2 空间的收敛性

- 逐点收敛

称函数列 f_n 在 $[a, b]$ 上逐点收敛到 f , 如果对于任意的 $t \in [a, b]$ 及任意的 $\epsilon > 0$, 存在正整数 N , 使得当 $n \geq N$ 时, $|f_n(t) - f(t)| < \epsilon$ 成立.

- 一致收敛

称函数列 f_n 在 $[a, b]$ 上一致收敛到 f , 如果对于任意的 $\epsilon > 0$, 存在正整数 N , 使得当 $n \geq N$ 时, 对任意的 $t \in [a, b]$, $|f_n(t) - f(t)| < \epsilon$ 成立.

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)

Page 23 of 66

[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

- 依范数收敛

称函数列 f_n 在 $L^2([a, b])$ 中依范数收敛到 f , 如果当 $n \rightarrow \infty$ 时, $\|f_n - f\|_{L^2} \rightarrow 0$. 更确切地说, 对任意的 $\epsilon > 0$, 存在正整数 N , 使得当 $n \geq N$ 时, $\|f_n - f\|_{L^2} < \epsilon$ 成立.

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [▶](#)

Page 24 of 66

[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

三种收敛的比较

- 逐点收敛与一致收敛

逐点收敛 \nRightarrow 一致收敛

一致收敛 \Rightarrow 逐点收敛

例 $f_n(t) = t^n, n = 1, 2, 3, \dots, f(t) = 0$.

❖ 在 $[0, 1)$ 上 f_n 逐点收敛到 f .

❖ 在 $[0, 1)$ 上 f_n 不一致收敛到 f . 因为当 t 越靠近 1 时, t^n 收敛于 0 的速度越慢.

❖ 对任意固定的 $r < 1$, 函数列 f_n 在 $[0, r]$ 上一致收敛于 f . 实际上, 如果 $0 \leq t \leq r$, 则 $|t^n| \leq r^n$. 于是只要选取 n 使得 $r^n < \epsilon$, 那么对任意的 $0 \leq t \leq r$, $|f_n(t) - f(t)| < \epsilon$ 成立.

Home Page

Title Page

◀ ▶

◀ ▶

Page 25 of 66

Go Back

Full Screen

Close

Quit

- 逐点收敛与依范数收敛

依范数收敛 \nRightarrow 逐点收敛

逐点收敛 \nRightarrow 依范数收敛

例1

$$f_n(t) = \begin{cases} \sqrt{n}, & 0 \leq t \leq \frac{1}{n^2}, \\ 0, & \frac{1}{n^2} < t \leq 1, \end{cases}$$
$$f(t) = 0, \quad t \in [0, 1].$$

$$\begin{aligned} \diamond \int_0^1 |f_n(t) - f(t)|^2 dt &= \int_0^{1/n^2} n dt = \frac{1}{n} \rightarrow 0. \\ \diamond f_n(0) &= \sqrt{n} \rightarrow \infty. \end{aligned}$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 26 of 66

Go Back

Full Screen

Close

Quit

例2 $f(t) = 0, t \in [0, 1]$,

$$f_n(t) = \begin{cases} 0, & t = 0, \\ n, & 0 < t < \frac{1}{n}, \\ 0, & \frac{1}{n} \leq t \leq 1. \end{cases}$$

❖ 函数列 f_n 在 $[0, 1]$ 上逐点收敛到 f . 因为对任意 $t \in [0, 1]$ 及 $\epsilon > 0$, 存在正整数 $N = [\frac{1}{t}]$, 使得当 $n \geq N$, 即 $n > \frac{1}{t}$ 时, $|f_n(t) - f(t)| = 0 < \epsilon$ 成立.

❖ 函数列 f_n 不依范数收敛到 f . 因为

$$\int_0^1 |f_n(t) - f(t)|^2 dt = \int_0^{\frac{1}{n}} n^2 dt = n \rightarrow \infty.$$

❖ 假设 $f_n, g \in L^2([a, b])$. 如果 f_n 在 $[a, b]$ 上逐点收敛到 f , 并且对任意的 $t \in [a, b]$, $|f_n(t)| \leq g(t)$, 则 f_n 在 $L^2[a, b]$ 中依范数收敛到 f .

Home Page

Title Page

◀ ▶

◀ ▶

Page 27 of 66

Go Back

Full Screen

Close

Quit

- 一致收敛与依范数收敛

依范数收敛 \nRightarrow 一致收敛

一致收敛 \nRightarrow 依范数收敛

例 $f(t) = 0, t \in [0, 1],$

$$f_n(t) = \begin{cases} 1, & 0 \leq t \leq \frac{1}{n}, \\ 0, & \frac{1}{n} < t \leq 1. \end{cases}$$

则 f_n 依范数收敛到 f , 但在 $[0, 1]$ 上 f_n 不一致收敛到 f . (Ex)

Home Page

Title Page

◀ ▶

◀ ▶

Page 28 of 66

Go Back

Full Screen

Close

Quit

定理: 如果在有限区间 $[a, b]$ 上, 函数列 f_n 一致收敛到 f , 在 $L^2([a, b])$ 中, f_n 依范数收敛于 f .

证明: 由一致收敛的定义可知, 对于任意的 $\epsilon > 0$, 存在正整数 N , 使得当 $n \geq N$ 时, 对任意的 $t \in [a, b]$, $|f_n(t) - f(t)| < \epsilon$ 成立. 于是

$$\begin{aligned}\|f_n - f\|_{L^2}^2 &= \int_a^b |f_n(t) - f(t)|^2 dt \\ &\leq \int_a^b \epsilon^2 dt \\ &= \epsilon^2(b - a).\end{aligned}$$

因此 f_n 依范数收敛于 f .

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [▶](#)

Page 29 of 66

[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

Schwarz 不等式与三角不等式

- 在欧式空间 \mathbb{R}^2 中, 由内积的定义可得 Schwarz 不等式

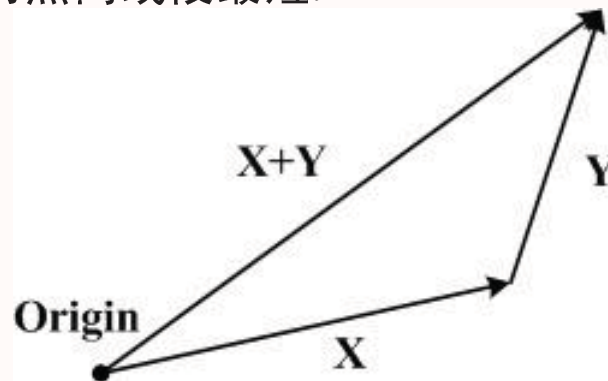
$$|\langle X, Y \rangle| = \|X\| \|Y\| |\cos \theta| \leq \|X\| \|Y\|,$$

其中 θ 是 X 和 Y 之间的夹角.

- 在欧式空间 \mathbb{R}^2 中, 三角不等式表现为

$$\|X + Y\| \leq \|X\| + \|Y\|.$$

该不等式描述为两点间线段最短.

[Home Page](#)[Title Page](#)[◀](#) [▶](#)[◀](#) [▶](#)[Page 30 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

定理: 假设 $V, \langle \cdot, \cdot \rangle$ 是一个内积空间, 则对任意的 $X, Y \in V$, 有

- **Schwarz 不等式:** $|\langle X, Y \rangle| \leq \|X\| \|Y\|$. 当且仅当 X 和 Y 线性相关时, 等号成立. 另外, 当且仅当 X 和 Y 是相互非负倍乘关系时, 有 $\langle X, Y \rangle = \|X\| \|Y\|$.
- **三角不等式:** $\|X + Y\| \leq \|X\| + \|Y\|$. 当且仅当 X 和 Y 是相互非负倍乘关系时, 等号成立.

Home Page

Title Page

◀ ▶

◀ ▶

Page 31 of 66

Go Back

Full Screen

Close

Quit

Schwarz 不等式的证明:

令 $\langle X, Y \rangle = |\langle X, Y \rangle|e^{i\phi}$, 并且 $t \in \mathbb{R}$. 考虑下面的不等式

$$\begin{aligned} 0 &\leq \|e^{-i\phi}X - tY\|^2 \\ &= \langle e^{-i\phi}X - tY, e^{-i\phi}X - tY \rangle \\ &= \|X\|^2 - t(\langle e^{-i\phi}X, Y \rangle + \langle Y, e^{-i\phi}X \rangle) + t^2\|Y\|^2 \\ &= \|X\|^2 - 2\operatorname{Re}(te^{-i\phi}\langle X, Y \rangle) + t^2\|Y\|^2 \\ &= \|X\|^2 - 2t|\langle X, Y \rangle| + t^2\|Y\|^2. \end{aligned}$$

于是, 判别式满足 $\Delta = 4|\langle X, Y \rangle|^2 - 4\|X\|^2\|Y\|^2 \leq 0$. 因此 Schwarz 不等式成立.

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 32 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

如果 $|\langle X, Y \rangle| = \|X\| \|Y\|$, 则方程

$$\|e^{-i\phi}X - tY\|^2 = 0$$

有双重实根 t_0 . 特别的, $e^{-i\phi}X - t_0Y = 0$, 即 $X = e^{i\phi}t_0Y$. 这表明 X 和 Y 线性相关. 如果 $\langle X, Y \rangle = \|X\| \|Y\|$, 则 $e^{i\phi} = 1$, 从而 $X = t_0Y$. 因此, 我们有

$$t_0\|Y\|^2 = |t_0|\|Y\|^2.$$

从而 $t_0 \geq 0$. 反之很容易证明.

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 33 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

三角不等式的证明:

根据 Schwarz 不等式, 可以得到

$$\begin{aligned}\|X + Y\|^2 &= \langle X + Y, X + Y \rangle \\ &= \|X\|^2 + 2 \operatorname{Re}(\langle X, Y \rangle) + \|Y\|^2 \\ &\leq \|X\|^2 + 2\|X\|\|Y\| + \|Y\|^2 \\ &= (\|X\| + \|Y\|)^2.\end{aligned}$$

不等式两边开平方即得三角不等式.

此外, 等号成立当且仅当 $\langle X, Y \rangle = \|X\|\|Y\|$, 因此等价于 X 和 Y 是相互非负倍乘关系.

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 34 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

平行四边形公式

定理: 假设 $V, \langle \cdot, \cdot \rangle$ 是一个内积空间, 则其诱导范数满足: 对任意的 $X, Y \in V$, 有

$$\|X + Y\|^2 + \|X - Y\|^2 = 2(\|X\|^2 + \|Y\|^2).$$

$$\begin{aligned} \text{证明: } \|X + Y\|^2 + \|X - Y\|^2 &= \langle X + Y, X + Y \rangle + \langle X - Y, X - Y \rangle \\ &= 2\langle X, X \rangle + 2\langle Y, Y \rangle \\ &= 2(\|X\|^2 + \|Y\|^2). \end{aligned}$$

在二维欧式空间中, 此公式表明平行四边形对角线长度的平方和等于四边长度的平方和.

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 35 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

正交

定义: 假设 $V, \langle \cdot, \cdot \rangle$ 是一个内积空间.

- 设 $X, Y \in V$, 如果 $\langle X, Y \rangle = 0$, 则称 X 和 Y 正交.
- 设 $X \in V$, M 是 V 的一个子集, 如果 X 与 M 中一切元素正交, 则称 X 与 M 正交.
- 设 M, N 是 V 的子集, 如果 M 中的任意元素都和 N 中的任意元素正交, 则称 M 与 N 正交.
- 设 M 是 V 的一个子集, 如果 M 中的元素都相互正交, 则称 M 是 V 的一个正交集. 进一步, 如果 M 中的元素还满足 $\|X\| = 1, \forall X \in M$, 则称 M 是 V 的一个标准正交集.

Home Page

Title Page

◀ ▶

◀ ▶

Page 36 of 66

Go Back

Full Screen

Close

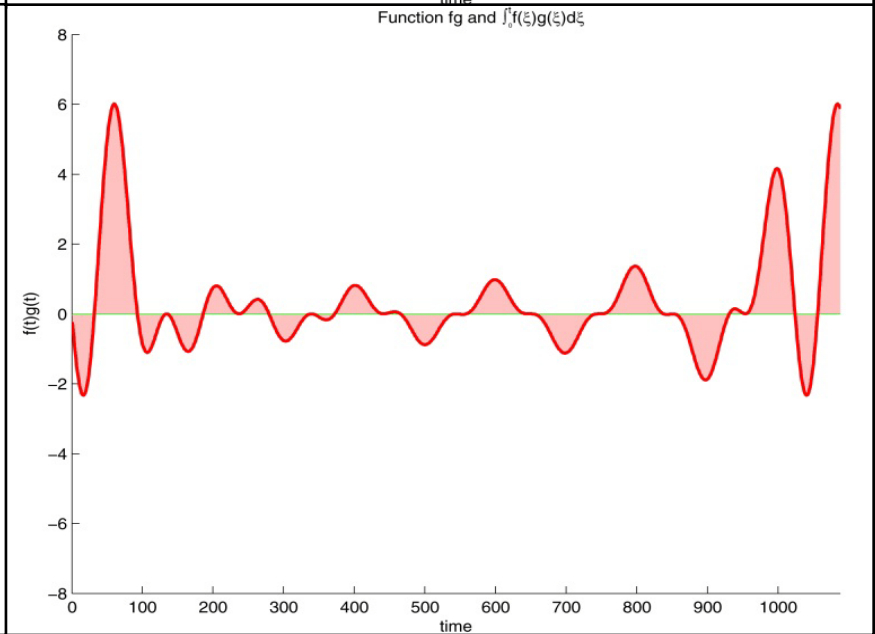
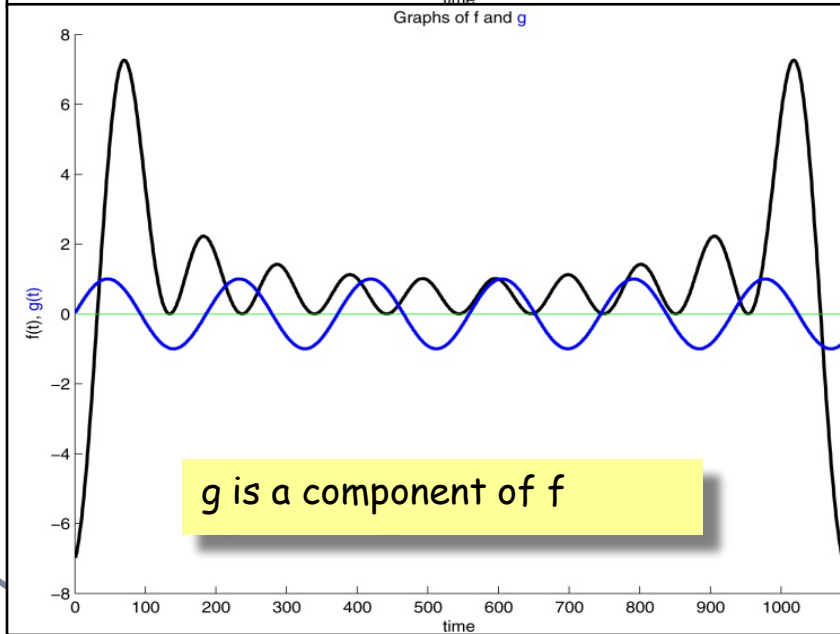
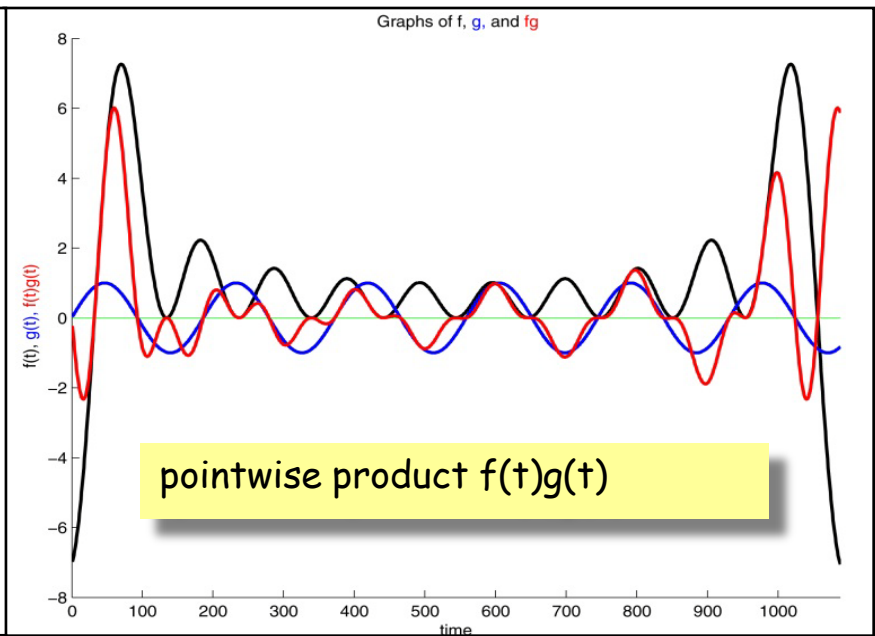
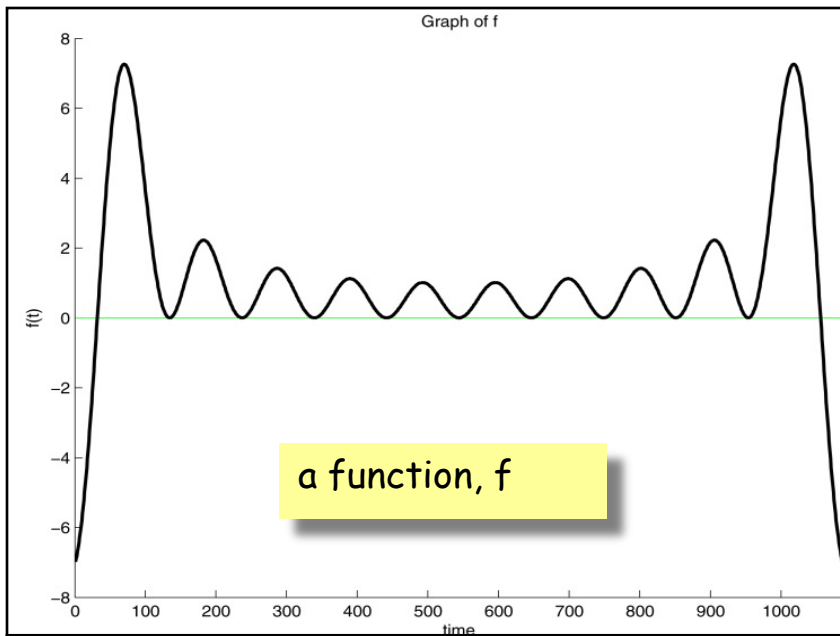
Quit

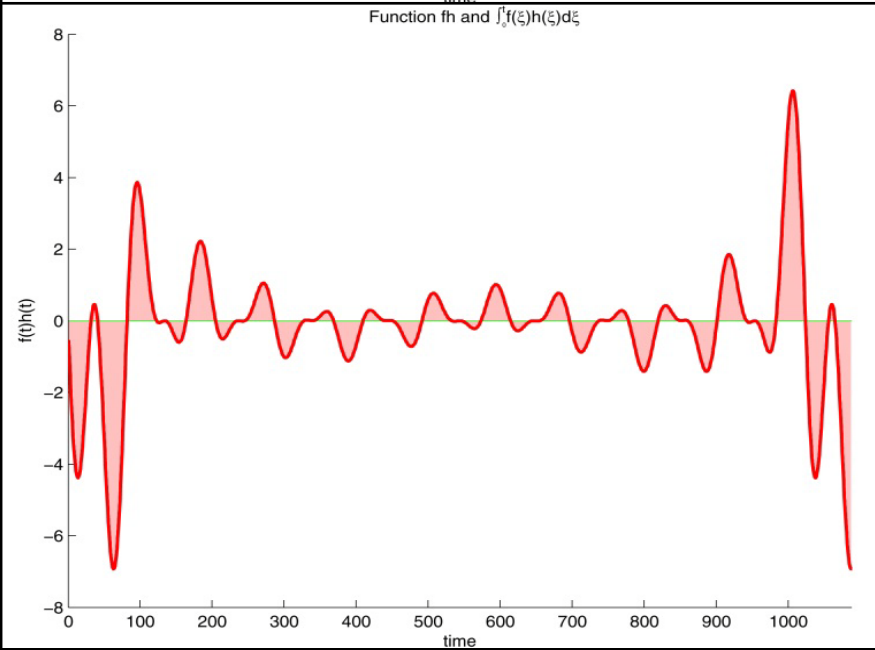
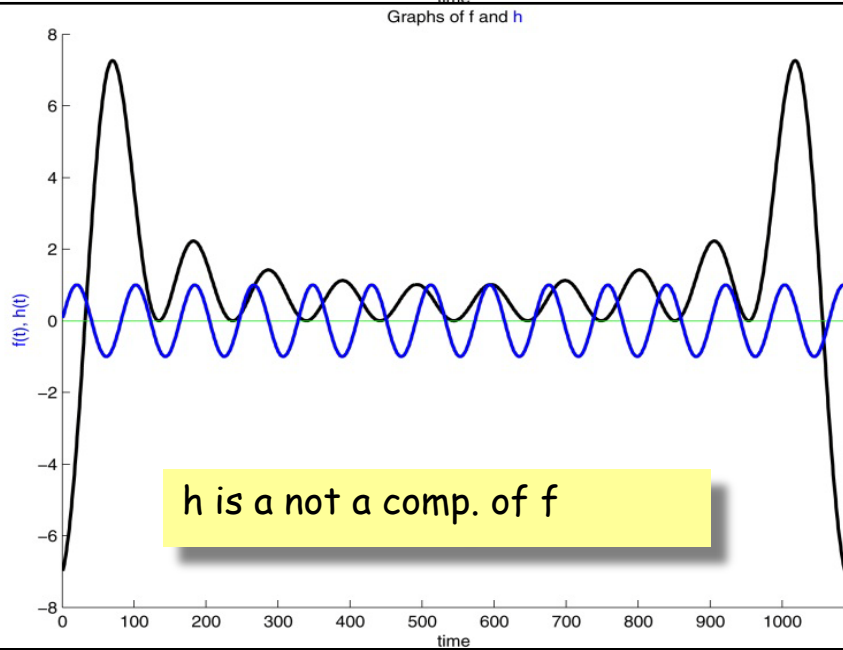
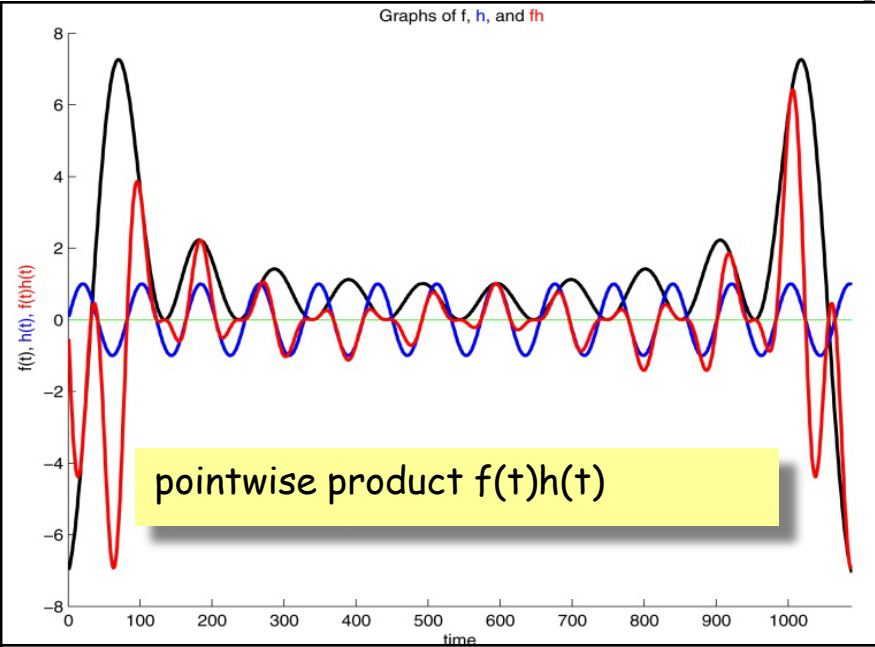
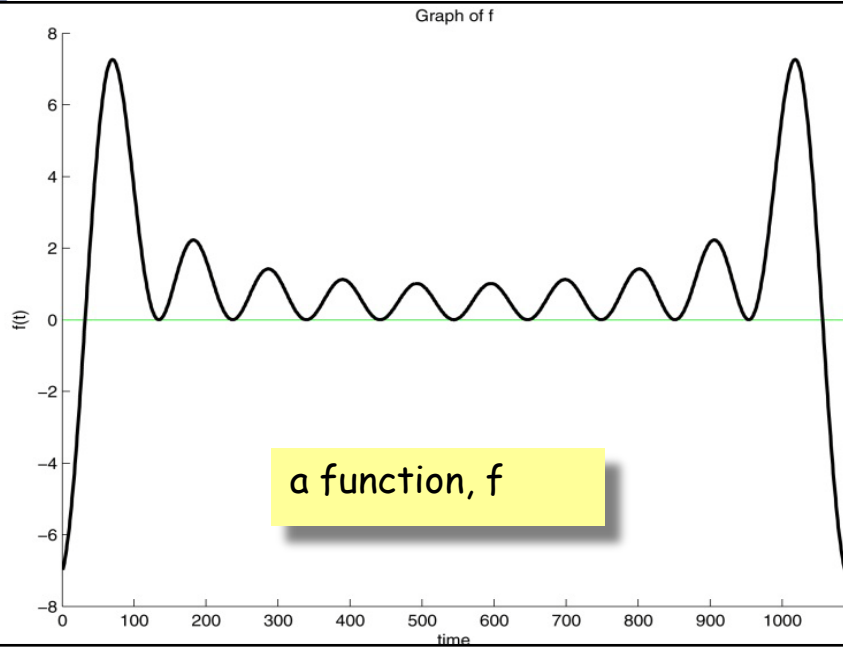
内积

内积

定义 1.4 空间 $L^2([a, b])$ 上的 L^2 内积定义为

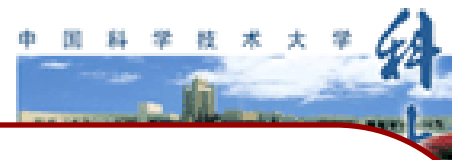
$$\langle f, g \rangle_{L^2} = \int_a^b f(t) \overline{g(t)} dt, f(t), g(t) \in L^2([a, b]).$$





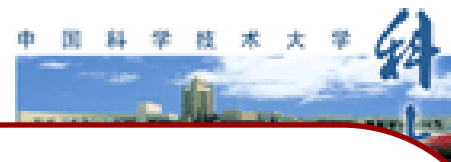
新闻分类

余弦定理和新闻分类



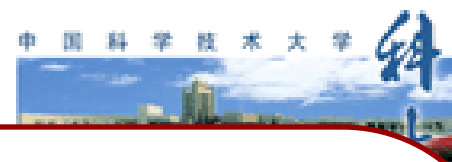
- 新闻的分类就是要将相似的新闻放到一类中；
- 计算机其实读不懂新闻，它只能快速计算；
- 这就要求我们设计一个算法来算出任意两篇新闻的相似性；
- 为了做到这一点，我们需要想办法用一组数字或一个向量来描述一篇新闻。

如何确定网页和查询的相关性



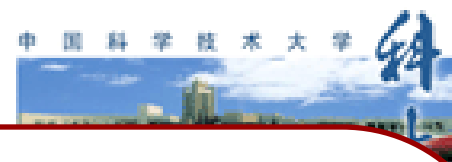
- 例子：查找关于“小波的应用”的网页
- 现在任何一个搜索引擎都包含几十万甚至是上百万个多少有点关系的网页。那么哪个应该排在前面呢？
- 显然我们应该根据网页和查询“小波的应用”的相关性对这些网页进行排序。
- 因此，这里的关键问题是如何度量网页和查询的相关性。

如何确定网页和查询的相关性



- 短语“小波的应用”可以分成三个关键词：“小波”、“的”、“应用”。
- 根据直觉，包含这三个词多的网页应该比包含它们少的网页相关。当然，这个办法有一个明显的漏洞。
- 因此我们需要根据网页的长度，对关键词的次数进行归一化，也就是用关键词的次数除以网页的总字数。我们把这个商称为“关键词的频率”，或者“单文本词汇频率”（Term Frequency, **TF**）

如何确定网页和查询的相关性



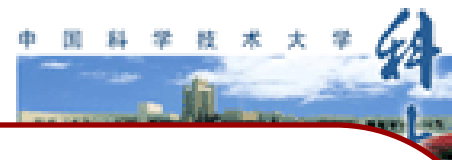
- 相关性的一个简单的度量:

如果一个查询包含关键词 w_1, w_2, \dots, w_n , 它们在一篇特定网页中的词频分别是: TF_1, TF_2, \dots, TF_n 。

(TF: term frequency)。那么, 这个查询和该网页的相关性就是:

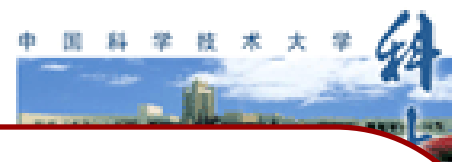
$$TF_1 + TF_2 + \dots + TF_n。$$

TF/IDF 介绍



- 在信息检索中，使用最多的权重是“逆文本频率指数”（Inverse document frequency 缩写为 **IDF**），
- 它的公式为 $\log(D / D_w)$ 其中 D 是全部网页数。比如，我们假定中文网页数是 $D = 10$ 亿，词“的”在所有的网页中都出现，即 $D_w = 10$ 亿，那么它的 $IDF = \log(10\text{亿}/10\text{亿}) = \log(1) = 0$ 。假如专用词“小波”在两百万个网页中出现，即 $D_w = 200$ 万，则它的权重 $IDF = \log(500) = 6.2$ 。又假定通用词“应用”，出现在五亿个网页中，它的权重 $IDF = \log(2)$ ，则只有 0.7。也就只说，在网页中找到一个“小波”的比配相当于找到九个“应用”的匹配。
- 利用 **IDF**，上述相关性计算公式就由词频的简单求和变成了加权求和，即 $TF_1 * IDF_1 + TF_2 * IDF_2 + \dots + TF_N * IDF_N$ 。

新闻的分类



- 怎样找一组数字，或者说一个向量来描述一篇新闻？
- 对于一篇新闻中的所有实词，我们可以计算出它们的单文本词汇频率/逆文本频率值（**TF/IDF**）。不难想象，和新闻主题有关的那些实词频率高，**TF/IDF** 值较大。

新闻的分类

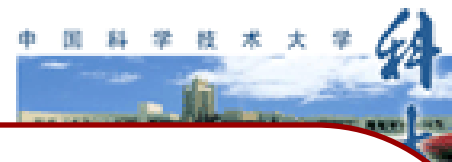


- 比如，词汇表有六万四千个词，分别为

单词编号 汉字词

1	阿
2	啊
3	阿斗
4	阿姨
...	
789	服装
....	
64000	做作

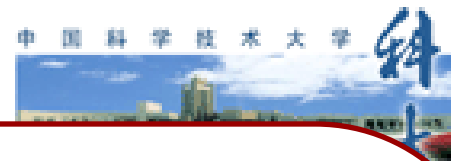
新闻的分类



- 在一篇新闻中，这 64,000 个词的 TF/IDF 值分别为

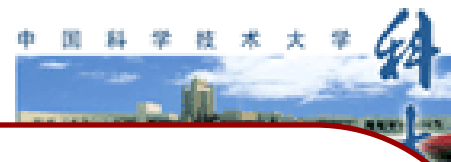
单词编号	TF/IDF 值
1	0
2	0.0034
3	0
4	0.00052
...	
789	0.034
...	
64000	0.075

新闻的分类



- 如果单词表中的某个词在新闻中没有出现，对应的值为零；
- 那么这 64,000 个数，组成一个64,000维的向量
- 我们就用这个向量来代表这篇新闻，并成为新闻的特征向量；
- 如果两篇新闻的特征向量相近，则对应的新闻内容相似，它们应当归在一类，反之亦然。

新闻的分类



- 向量实际上是多维空间中有方向的线段。如果两个向量的方向一致，即夹角接近零，那么这两个向量就相近。
- 而要确定两个向量方向是否一致，这就要用到余弦定理计算向量的夹角了。

新闻的分类



- 三角形的余弦定理公式:

$$\cos A = \frac{b^2 + c^2 - a^2}{2bc}$$

- 将三角形的两边 **b** 和 **c** 看成是两个向量，那么上述公式等价于

其中分母表示两个向量 **b** 和 **c** 的长度，分子表示两个向量的内积

$$\cos A = \frac{\langle \mathbf{b}, \mathbf{c} \rangle}{|\mathbf{b}| |\mathbf{c}|}$$

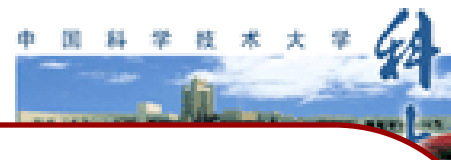
新闻的分类



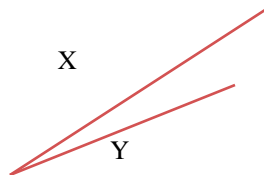
举一个具体的例子，假如新闻 X 和新闻 Y 对应向量分别是 $x_1, x_2, \dots, x_{64000}$ 和 $y_1, y_2, \dots, y_{64000}$ ，那么它们夹角的余弦等于：

$$\cos \theta = \frac{x_1 y_1 + x_2 y_2 + \dots + x_{64000} y_{64000}}{\sqrt{x_1^2 + x_2^2 + \dots + x_{64000}^2} \cdot \sqrt{y_1^2 + y_2^2 + \dots + y_{64000}^2}}$$

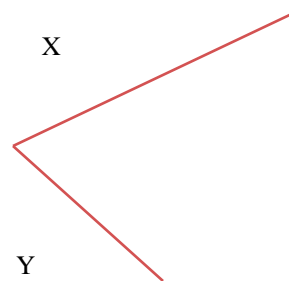
新闻的分类



- 当两条新闻向量夹角的余弦等于一时，这两条新闻完全重复（用这个办法可以删除重复的网页）；
- 当夹角的余弦接近于一时，两条新闻相似，从而可以归成一类；
- 夹角的余弦越小，两条新闻越不相关。



两天新闻相似



两天新闻无关

正交

定义: 假设 $V, \langle \cdot, \cdot \rangle$ 是一个内积空间.

- 设 $X, Y \in V$, 如果 $\langle X, Y \rangle = 0$, 则称 X 和 Y 正交.
- 设 $X \in V$, M 是 V 的一个子集, 如果 X 与 M 中一切元素正交, 则称 X 与 M 正交.
- 设 M, N 是 V 的子集, 如果 M 中的任意元素都和 N 中的任意元素正交, 则称 M 与 N 正交.
- 设 M 是 V 的一个子集, 如果 M 中的元素都相互正交, 则称 M 是 V 的一个正交集. 进一步, 如果 M 中的元素还满足 $\|X\| = 1, \forall X \in M$, 则称 M 是 V 的一个标准正交集.

Home Page

Title Page

◀ ▶

◀ ▶

Page 36 of 66

Go Back

Full Screen

Close

Quit

例 1 令

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1, \\ 0, & \text{otherwise,} \end{cases} \quad \psi(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq t < 1, \\ 0, & \text{otherwise.} \end{cases}$$

则在 $L^2(\mathbb{R})$ 中, ϕ 和 ψ 是正交的, 这是因为

$$\langle \phi, \psi \rangle = \int_0^{\frac{1}{2}} 1 dt - \int_{\frac{1}{2}}^1 1 dt = 0.$$

◆ 在 Haar 系里, ϕ 称为尺度函数([scaling function](#)), ψ 称为小波函数([wavelet function](#)).

[Home Page](#)

[Title Page](#)

[◀◀](#) [▶▶](#)

[◀](#) [▶](#)

Page 37 of 66

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

例 2 函数 $f(t) = \sin t$ 和 $g(t) = \cos t$ 在 $L^2([-\pi, \pi])$ 中正交, 这是因为

$$\langle f, g \rangle = \int_{-\pi}^{\pi} \sin t \cos t dt = \frac{1}{2} \int_{-\pi}^{\pi} \sin 2t dt = 0.$$

此外, 由

$$\int_{-\pi}^{\pi} \sin^2 t dt = \int_{-\pi}^{\pi} \cos^2 t dt = \pi,$$

可知 $\frac{\sin t}{\sqrt{\pi}}$ 和 $\frac{\cos t}{\sqrt{\pi}}$ 在 $L^2([-\pi, \pi])$ 中是标准正交的.

更一般的, 函数系

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos nt}{\sqrt{\pi}}, \frac{\sin nt}{\sqrt{\pi}}, n = 1, 2, \dots$$

在 $L^2([-\pi, \pi])$ 中是标准正交的. 该函数系在 Fourier 级数理论中是非常重要的.

Home Page

Title Page

◀ ▶

◀ ▶

Page 38 of 66

Go Back

Full Screen

Close

Quit

标准正交基

定义 内积空间的向量基如果是标准正交的, 则称其为内积空间的标准正交基.

定理 假设 V_0 是内积空间 V 的一个子空间, $\{e_1, \dots, e_N\}$ 是 V_0 的标准正交基. 如果 $v \in V_0$, 则有

$$v = \sum_{j=1}^N \langle v, e_j \rangle e_j.$$

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 39 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

证明 因为 $\{e_1, \dots, e_N\}$ 是 V_0 的标准正交基, 所以任意向量 $v \in V_0$ 可以唯一的表示成为 e_j 的线性组合, 即

$$v = \sum_{j=1}^N \alpha_j e_j.$$

上式两端对 e_k 取内积, 则有

$$\langle v, e_k \rangle = \sum_{j=1}^N \langle \alpha_j e_j, e_k \rangle.$$

由 $\{e_1, \dots, e_N\}$ 的标准正交性可得

$$\alpha_k = \langle v, e_k \rangle.$$

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 40 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

正交投影与逼近

最佳逼近问题 假设 V 是一个内积空间, V_0 是 V 的一个子空间, 并且 $v \in V \setminus V_0$. 很自然的提出两个问题:

① 是否存在 $v_0 \in V_0$, 使得 v_0 是 v 到 V_0 的最佳逼近? 也就是说,

$$\|v - v_0\| = \min_{w \in V_0} \|v - w\|. \quad \text{💬}$$

② v 到 V_0 的最佳逼近如果存在是否唯一?

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 41 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

定理 假设 V 是一个内积空间, V_0 是 V 的完备子空间. 则对任意的 $v \in V$, 存在唯一的 $v_0 \in V_0$, 使得

$$\|v - v_0\| = \min_{w \in V_0} \|v - w\|.$$

称 v_0 是 v 到 V_0 的正交投影.

◆ 特别的, 如果 V_0 是 V 的有限维子空间, 则是 V 的完备子空间, 因此上述结论成立.

Home Page

Title Page

◀ ▶

◀ ▶

Page 42 of 66

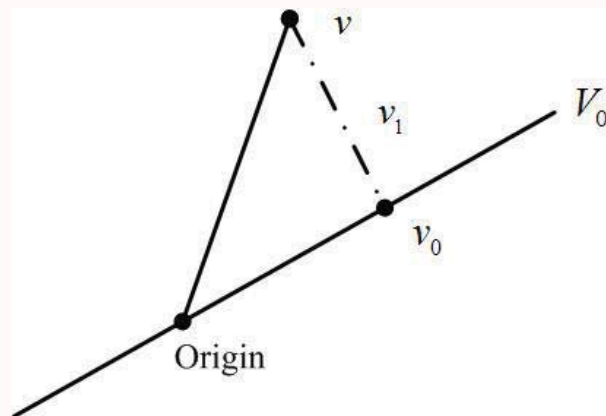
Go Back

Full Screen

Close

Quit

投影定理 假设 V 是一个内积空间, V_0 是 V 的有限维子空间, 并且 v 是 V 中任一元素. 则 $v_0 \in V_0$ 是 v 到 V_0 的正交投影当且仅当 $v - v_0$ 与 V_0 正交.

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 43 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

证明 (\Rightarrow) 对任意的 $w \in V_0$, 考虑函数

$$f(t) = \|v_0 + tw - v\|^2, \quad t \in \mathbb{R}.$$

因为 v_0 是 v 到 V_0 的正交投影, 所以 $t = 0$ 是 f 的最小值点. 利用内积将 f 表示如下

$$\begin{aligned} f(t) &= \langle v_0 + tw - v, v_0 + tw - v \rangle \\ &= \|v_0 - v\|^2 + 2t \operatorname{Re}(\langle v_0 - v, w \rangle) + t^2 \|w\|^2, \end{aligned}$$

于是可得

$$f'(t) = 2 \operatorname{Re}(\langle v_0 - v, w \rangle) + 2t \|w\|^2.$$

因为 $t = 0$ 是 f 的最小值点, 所以 $f'(0) = 0$. 于是我们得到

$$\operatorname{Re}(\langle v_0 - v, w \rangle) = 0.$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 44 of 66

Go Back

Full Screen

Close

Quit

同理, 考虑函数

$$g(t) = \|v_0 + itw - v\|^2, t \in \mathbb{R}.$$

因为 v_0 是 v 到 V_0 的正交投影, 所以 $t = 0$ 也是 g 的最小值点. 将 g 展开

$$\begin{aligned} g(t) &= \langle v_0 + itw - v, v_0 + itw - v \rangle \\ &= \|v_0 - v\|^2 + 2t \operatorname{Im}(\langle v_0 - v, w \rangle) + t^2 \|w\|^2, \end{aligned}$$

可得

$$g'(t) = 2 \operatorname{Im}(\langle v_0 - v, w \rangle) + 2t \|w\|^2.$$

由 $g'(0) = 0$, 可知 $\operatorname{Im}(\langle v_0 - v, w \rangle) = 0$. 从而得到结论

$$\langle v_0 - v, w \rangle = 0, \forall w \in V_0.$$

Home Page

Title Page

◀ ▶

◀ ▶

Page 45 of 66

Go Back

Full Screen

Close

Quit

(\Leftarrow) 假设 $v_0 \in V_0, v - v_0 \perp V_0$, 并且 $\tilde{v}_0 \in V_0$ 是 v 到 V_0 的正交投影.
于是根据前面的论证可知,

$$v - \tilde{v}_0 \perp V_0.$$

又因为 $\tilde{v}_0 - v_0 \in V_0$, 并且

$$\tilde{v}_0 - v_0 = (v - v_0) - (v - \tilde{v}_0)$$

和 V_0 正交, 从而 $\tilde{v}_0 - v_0 = 0$, 即 $\tilde{v}_0 = v_0$.

Home Page

Title Page

◀ ▶

◀ ▶

Page 46 of 66

Go Back

Full Screen

Close

Quit

正交投影的表示

定理 假设 V 是一个内积空间, V_0 是 V 的 N 维子空间, 并且 $\{e_1, e_2, \dots, e_N\}$ 是 V_0 的标准正交基. 则对任意的 $v \in V$, 其到 V_0 的正交投影表示为

$$v_0 = \sum_{j=1}^N \alpha_j e_j, \quad \alpha_j = \langle v, e_j \rangle.$$

证明 因为 $\{e_1, e_2, \dots, e_N\}$ 是 V_0 的基底, 所以要证明 $v - v_0$ 正交于 V_0 , 只需说明 $v - v_0$ 与 e_k , $k = 1, \dots, N$ 正交. 这可由下面得到

$$\begin{aligned} \langle v - v_0, e_k \rangle &= \langle v, e_k \rangle - \left\langle \sum_{j=1}^N \alpha_j e_j, e_k \right\rangle \\ &= \langle v, e_k \rangle - \alpha_k \\ &= 0. \end{aligned}$$

[Home Page](#)[Title Page](#)[◀](#) [▶](#)[◀](#) [▶](#)[Page 47 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

例1 令 V_0 是 $L^2([-\pi, \pi])$ 中由 $\cos t$ 和 $\sin t$ 张成的子空间, $f(t) = t$. 因为 $e_1 = \frac{\cos t}{\sqrt{\pi}}$ 和 $e_2 = \frac{\sin t}{\sqrt{\pi}}$ 在 $L^2([-\pi, \pi])$ 中是标准正交的, 所以 f 到 V_0 的正交投影为

$$f_0 = \langle f, e_1 \rangle e_1 + \langle f, e_2 \rangle e_2,$$

其中

$$\langle f, e_1 \rangle = 0, \quad \langle f, e_2 \rangle = 2\sqrt{\pi}.$$

于是可得

$$f_0(t) = 2 \sin t.$$

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 48 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

例2 令

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1, \\ 0, & \text{otherwise}, \end{cases} \quad \psi(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq t < 1, \\ 0, & \text{otherwise}. \end{cases}$$

设 V 是 $L^2[0, 1]$ 中由 $\phi(t), \psi(t), \psi(2t), \psi(2t - 1)$ 张成的子空间. 求 $f(t) = t$ 到 V 的正交投影.(Ex)

Home Page

Title Page

◀ ▶

◀ ▶

Page 49 of 66

Go Back

Full Screen

Close

Quit

内积空间的正交分解

正交补空间 假设 V 是一个内积空间, V_0 是 V 的一个子空间. V_0 的正交补空间, 记为 V_0^\perp , 是 V 中所有与 V_0 正交的向量的集合, 即

$$V_0^\perp = \{v \in V; \langle v, w \rangle = 0, \forall w \in V_0\}.$$

正交分解定理 假设 V 是一个内积空间, V_0 是 V 的有限维子空间. 则 V 中任一元素 v 能够被唯一表示成

$$v = v_0 + v_1,$$

其中 $v_0 \in V_0, v_1 \in V_0^\perp$. 此即表明

$$V = V_0 \oplus V_0^\perp.$$

[Home Page](#)[Title Page](#)[◀◀](#) [▶▶](#)[◀](#) [▶](#)[Page 50 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

证明 对任意的 $v \in V$, 令 v_0 是 v 到 V_0 的正交投影. 于是 v 可以表示成

$$v = v_0 + (v - v_0),$$

并且 $v - v_0$ 与 V_0 正交, 也就是说 $v - v_0 \in V_0^\perp$.

下面只需要给出分解的唯一性. 假设有两种分解形式

$$v = v_0 + v_1 = \tilde{v}_0 + \tilde{v}_1,$$

其中 $v_0, \tilde{v}_0 \in V_0, v_1, \tilde{v}_1 \in V_0^\perp$. 于是我们有

$$V_0 \ni v_0 - \tilde{v}_0 = \tilde{v}_1 - v_1 \in V_0^\perp.$$

这表明

$$\langle v_0 - \tilde{v}_0, v_0 - \tilde{v}_0 \rangle = \langle \tilde{v}_1 - v_1, \tilde{v}_1 - v_1 \rangle = 0.$$

从而可得 $v_0 = \tilde{v}_0, v_1 = \tilde{v}_1$.

Home Page

Title Page

◀ ▶

◀ ▶

Page 51 of 66

Go Back

Full Screen

Close

Quit

Gram-Schmidt 正交化

定理 假设 V 是一个内积空间, V_0 是 V 的 N 维子空间, 并且 $v_j, j = 1, \dots, N$, 是 V_0 的一组基. 则存在 V_0 的标准正交基 $\{e_1, e_2, \dots, e_N\}$, 满足每一个 e_j 都是 v_1, \dots, v_j 的线性组合.

[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 52 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

证明: 首先定义 $e_1 := \frac{v_1}{\|v_1\|}$. 很明显, $\|e_1\| = 1$. 令 v_0 是 v_2 到子空间 $\text{span}\{e_1\}$ 的正交投影. 则

$$v_0 = \langle v_2, e_1 \rangle e_1,$$

并且向量

$$E_2 = v_2 - v_0 = v_2 - \langle v_2, e_1 \rangle e_1$$

与 e_1 正交. 注意到 $E_2 \neq 0$, 否则 v_2 和 v_1 是线性相关的. 于是定义 $e_2 := \frac{E_2}{\|E_2\|}$. 显然, e_1 和 e_2 是标准正交的, 并且 e_2 是 v_1 和 v_2 的线性组合.

[Home Page](#)[Title Page](#)[<<](#) [>>](#)[<](#) [>](#)[Page 53 of 66](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

令 v_0 是 v_3 到子空间 $\text{span}\{e_1, e_2\}$ 的正交投影. 则

$$v_0 = \langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2,$$

并且向量 $E_3 = v_3 - v_0$ 与 e_1 和 e_2 正交. 定义 $e_3 := \frac{E_3}{\|E_3\|}$. 同理可得 $\{e_1, e_2, e_3\}$ 是标准正交集, 并且 e_3 可表示成 v_1, v_2, v_3 的线性组合.

继续此过程, 可以得到定理结论.

例: 设 V 是 $L^2[0, 1]$ 中由 $1, t, t^2, t^3$ 张成的子空间. 利用 Gram-Schmidt 正交化方法构造 V 的标准正交基.(Ex)

Home Page

Title Page

◀ ▶

◀ ▶

Page 54 of 66

Go Back

Full Screen

Close

Quit

正交化

设 $v_1(t), v_2(t), \dots, v_n(t)$ 是 $L^2[a, b]$ 中一组线性无关函数系, 记 $\psi_1(t) = v_1(t)$,

$$\psi_n(t) = \begin{vmatrix} \langle v_1, v_1 \rangle, \dots, \langle v_1, v_{n-1} \rangle & v_1 \\ \langle v_2, v_1 \rangle, \dots, \langle v_2, v_{n-1} \rangle & v_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \vdots \\ \langle v_n, v_1 \rangle, \dots, \langle v_n, v_{n-1} \rangle & v_n \end{vmatrix}$$

以及 $\Delta_0 := 1, \Delta_n := \det(\langle v_i, v_j \rangle)_{i,j=1}^n$, 则

$$\varphi_k(t) = \frac{\psi_k(t)}{\sqrt{\Delta_k \Delta_{k-1}}}, k \geq 1 \text{ 是一组规范正交系, 且}$$

$$\text{span}\{\varphi_i\}_1^k = \text{span}\{v_i\}_1^k, k = 1, 2, \dots, n \quad (1.4)$$

证明 由 $\psi_n(t)$ 的定义, 显然有 $\langle \psi_n, v_k \rangle = \begin{cases} \Delta_n & k = n \\ 0 & k < n \end{cases}$, 因而, 当 $k \leq n$ 时,

$$\begin{aligned} \langle \varphi_n, \varphi_k \rangle &= \frac{\langle \psi_n, \psi_k \rangle}{\sqrt{\Delta_n \Delta_{n-1} \Delta_k \Delta_{k-1}}} \\ &= \frac{1}{\sqrt{\Delta_n \Delta_{n-1} \Delta_k \Delta_{k-1}}} \left\langle \psi_n, \sum_{i=1}^{k-1} \beta_i v_i + \frac{\Delta_{k-1}}{\sqrt{\Delta_k \Delta_{k-1}}} v_k \right\rangle \\ &= \delta_{k,n}. \end{aligned} \quad (1.5)$$

这就证明了 $\{\varphi_k\}_{k=1}^n$ 的标准正交性, 又因为

$$\varphi_{k+1} = \frac{\Delta_k}{\sqrt{\Delta_{k+1} \Delta_k}} v_{k+1} + \sum_{i=1}^k \alpha_i v_i \in M_{k+1}$$

但 $\varphi_{k+1} \notin \text{span}\{v_i\}_{i=1}^k$, 所以, $\text{span}\{\varphi_i\}_{i=1}^{k+1} = \text{span}\{v_i\}_{i=1}^{k+1}$.

#

解矛盾方程组



■ 矛盾方程组:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m \times n}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n \times 1}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}_{m \times 1}$$

■ 当 $m \gg n$ 时, 方程组 $\mathbf{Ax} = \mathbf{b}$ 通常无解, 故称为矛盾方程组

■ 修改: 要求均方误差最小

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2^2$$

- 最小二乘意义下的解

解矛盾方程组



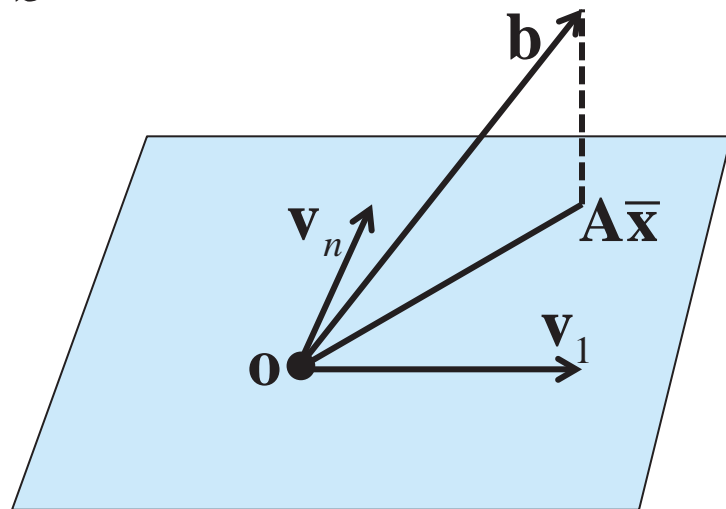
■ 最小二乘法的几何解释:

$$\mathbf{Ax} = \mathbf{b} \Rightarrow x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n = \mathbf{b}$$

$$(\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) \perp \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n\}$$

$$\mathbf{x}^T \mathbf{A}^T (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) = 0, \quad \forall \mathbf{x} \in \mathbb{R}^n$$

$$\mathbf{A}^T (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) = 0 \Rightarrow \mathbf{A}^T \mathbf{A}\bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$$



解矛盾方程组



■ 定理：设 A 为 $m \times n$ 阶的矩阵， b 为 m 维列向量，则

(1) $A^T A x = A^T b$ 称为矛盾方程 $Ax = b$ 的法方程，恒有解

(2) x 是 $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ 的解 $\Leftrightarrow x$ 是法方程 $A^T A x = A^T b$ 的解

线性拟合和二次拟合函数



- 最小二乘问题：给定数据序列 $\{(x_i, y_i)\}_{i=1}^m$ 和函数空间

$$\Phi = \text{span}\{\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)\},$$

求函数 $\varphi(x) \in \Phi$ ，使得

$$\min_{\varphi \in \Phi} \sum_{i=1}^m (\varphi(x_i) - y_i)^2$$

其中 $\varphi(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \dots + a_n\varphi_n(x)$, $a_i \in \mathbb{R}$, $i = 1, 2, \dots, n$

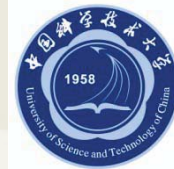
- 若定义

$$Q(a_1, a_2, \dots, a_n) = \sum_{i=1}^m (\varphi(x_i) - y_i)^2 = \sum_{i=1}^m (a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + \dots + a_n\varphi_n(x_i) - y_i)^2$$

则最小二乘问题可写成

$$\min_{a_1, a_2, \dots, a_n \in \mathbb{R}} Q(a_1, a_2, \dots, a_n)$$

线性拟合和二次拟合函数



- 将 $Q(a_1, a_2, \dots, a_n) = \sum_{i=1}^m (a_1 \varphi_1(x_i) + a_2 \varphi_2(x_i) + \dots + a_n \varphi_n(x_i) - y_i)^2$ 写成矩阵的表示形式：记

$$\mathbf{A} = \begin{pmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m) \end{pmatrix}_{m \times n}, \quad \mathbf{x} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}_{n \times 1}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}_{m \times 1},$$

则有

$$Q(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2 = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \mathbf{b}^T \mathbf{b}$$

- 二次型取最小值的条件

$$\frac{\partial Q(\mathbf{x})}{\partial \mathbf{x}} = 0 \Rightarrow 2\mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{A}^T \mathbf{b} = 0 \Rightarrow \mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}$$

线性拟合和二次拟合函数



- 线性拟合：取函数空间 $\Phi = \text{span}\{1, x\}$ ， $\varphi(x) = a + bx$ ，则

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b} \quad \longrightarrow \quad \begin{pmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{pmatrix}$$

线性拟合和二次拟合函数



- 二次拟合：取函数空间 $\Phi = \text{span}\{1, x, x^2\}$, $\varphi(x) = a + bx + cx^2$, 则

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b} \quad \longrightarrow \quad \begin{pmatrix} m & \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 \\ \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 & \sum_{i=1}^m x_i^4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m x_i^2 y_i \end{pmatrix}$$

线性拟合和二次拟合函数



■ 形如 ae^{bx} 曲线拟合：取函数空间 $\Phi = \{ae^{bx} \mid a, b \in \mathbb{R}\}$

- 非线性空间
- 线性化：作变换 $z = \ln y$ ，有

$$z_i = \ln y_i$$

$$\ln \varphi(x) = \ln(ae^{bx}) = \ln a + bx = A + Bx$$

- 利用线性拟合求 A, B

■ 类似地，可求形如 $\frac{1}{a+bx}$ 曲线拟合

■ 注意：变换后的所得拟合曲线，已经不在是平方误差极小意义下的拟合曲线

数值积分



- Newton-Leibniz公式, Green公式, Gauss公式, Stokes公式
- 很多积分无解析解, 必须使用数值方式求解
- 从积分的定义出发

$$I(f) = \int_a^b f(x) dx \approx \lim_{\Delta x_i \rightarrow 0} \left(\sum_{i=0}^n f(x_i) \Delta x_i \right)$$

$$\approx \sum_{i=0}^n \alpha_i f(x_i) = I_n(f)$$

- 求积节点 $\{x_i\}$
- 求积系数 $\{\alpha_i\}$

■ 代数精度：衡量数值积分公式优劣的重要指标之一

- 设 $[a, b]$ 上以 $x_i, i = 0, 1, \dots, n$ 为积分节点的数值积分公式为

$$I_n(f) = \sum_{i=0}^n \alpha_i f(x_i),$$

若 $I_n(f)$ 满足

$$I_n(x^i) = I(x^i), i = 0, \dots, k;$$

$$I_n(x^{k+1}) \neq I(x^{k+1}),$$

则称 $I_n(f)$ 具有 k 阶代数精度

- ## ■ 性质：当 $I_n(f)$ 具有 k 阶代数精度时，对任意不高于 k 次的多项式 $p(x)$ ，有 $I_n(p) = I(p)$

■ 插值型数值积分

- 思想：用插值函数，譬如Lagrange插值函数，代替被积函数，进行积分

$$\begin{aligned} I(f) \approx I_n(f) &= \int_a^b L_n(x) dx = \int_a^b \sum_{i=0}^n l_i(x) f(x_i) dx = \sum_{i=0}^n \left(\int_a^b l_i(x) dx \right) f(x_i) \\ &= \sum_{i=0}^n \alpha_i f(x_i) \end{aligned}$$

$$\text{其中 } \alpha_i = \int_a^b l_i(x) dx$$

■ 误差估计公式

$$E_n(f) = I(f) - I_n(f) = \int_a^b R_n(x) dx = \int_a^b \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \omega_n(x) dx$$

■ 代数精度

- n 次插值多项式形式的数值积分公式至少有 n 阶代数精度

Gauss型积分



- 对于Newton-Cotes积分, 奇数有 n 阶代数精度, 偶数有 $n+1$ 阶精度
- 如果放弃等距节点的要求, 是否能建立更有高精度的公式?
- 例: 取区间 $[-1, 1]$, 考虑两点数值积分格式, 将积分点视为未知量, 则有

$$\left\{ \begin{array}{l} a_0 \cdot 1 + a_1 \cdot 1 = \int_{-1}^1 1 dx = 2 \\ a_0 \cdot x_0 + a_1 \cdot x_1 = \int_{-1}^1 x dx = 0 \\ a_0 \cdot x_0^2 + a_1 \cdot x_1^2 = \int_{-1}^1 x^2 dx = \frac{2}{3} \\ a_0 \cdot x_0^3 + a_1 \cdot x_1^3 = \int_{-1}^1 x^3 dx = 0 \end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{l} a_0 = 1 \\ a_1 = 1 \\ x_0 = -\frac{1}{\sqrt{3}} \\ x_1 = \frac{1}{\sqrt{3}} \end{array} \right.$$

不难验证数值积分公式 $\int_{-1}^1 f(x) dx = f(-\frac{1}{\sqrt{3}}) + f(\frac{1}{\sqrt{3}})$ 具有三阶代数精度

Gauss型积分



■ **定理：** 设 $I(f) = \int_a^b f(x)dx$ 关于积分节点 x_1, x_2, \dots, x_n 的数值积分公式为 $I_n(f) = \sum_{i=1}^n a_i f(x_i)$ ，则 $I_n(f)$ 的代数精度不超过 $2n-1$ 阶

证明： 取多项式 $p(x) = (x-x_1)^2(x-x_2)^2 \cdots (x-x_n)^2 = \omega_n^2(x)$

易知 $I(p(x)) > 0$ 及 $I_n(p(x)) = 0$

故数值积分公式 $I_n(f)$ 的代数精度不可能达到 $2n$ 阶

■ **如何构造最高阶 $(2n-1)$ 精度的公式？**

■ **Gauss型积分：** 取一组特殊的积分节点（正交函数的零点）

Gauss型积分



- 定义：给定一般形式的积分和内积

$$I(f) = \int_a^b W(x) f(x) dx, \quad W(x) \geq 0,$$

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

其中 $W(x)$ 为权函数. 若 $\langle f, g \rangle = 0$, 则称函数 f 与 g 正交

- 如何将线性函数空间的一组基变为一组正交基?
- 利用Schmidt正交化过程:

$$\begin{cases} g_1(x) = f_1(x) \\ \vdots \\ g_n(x) = f_n(x) - \sum_{i=1}^{n-1} \frac{(f_n(x), g_i(x))}{(g_i(x), g_i(x))} g_i(x) \end{cases}$$

Gauss型积分



■ **定理：**以 n 次正交多项式的 n 个零点为积分的数值积分公式有 $2n-1$ 阶代数精度

■ **证明：**设 n 次正交多项式 $p_n(x)$ 的 n 个零点为 x_1, x_2, \dots, x_n ，
记 $G_n(f) = \int_a^b L_n(x)W(x)dx = \sum_{i=1}^n \left(\int_a^b l_i(x)W(x)dx \right) f(x_i) = \sum_{i=1}^n \alpha_i f(x_i)$ ，
则有

$$E(f) = I(f) - G_n(f) = \int_a^b f[x_1, x_2, \dots, x_n, x] \omega_n(x) W(x) dx$$

对于任意的 $p(x) \in P_{2n-1}(x)$ ，利用多项式带余除法得

$$p(x) = s(x)p_n(x) + r(x), \quad s(x), r(x) \in P_{n-1}(x),$$

$$\text{故 } \int_a^b p(x)W(x)dx = \int_a^b [s(x)p_n(x) + r(x)]W(x)dx = \int_a^b r(x)W(x)dx.$$

$$\text{因此 } E(p) = I(p) - G_n(p) = I(r) - G_n(r) = E(r) = 0$$

■ **收敛性：**若 $f(x) \in C[a, b]$ ，则 $\lim_{n \rightarrow \infty} G_n(f) = I(f)$

仅理论上可行，因为求 n 次正交多项式的所有零点是困难的！

作业

10. 计算向量 $(1, 2, 1)$ 张成的 R^3 子空间的正交补空间。
11. 如果 $f(t) = 1, 0 \leq t \leq 1$, 则在 $L^2[0, 1]$ 上 $f(t)$ 的正交补空间是均值为0的所有函数组成的空间。
12. 利用Gram-Schmidt正交化方法求由 $\{1, x, x^2\}$ 张成的 $L^2[0, 1]$ 的子空间的标准正交基。
13. 求 $\sin(x) + \cos(x)$ 在由 $\{1, x, x^2\}$ 张成的 $L^2[0, 1]$ 的子空间的正交投影。
14. 设 e_1, \dots, e_k 是 n 维内积空间 V 的一组两两正交的单位向量, 给定 $v \in V$, 记 $\alpha_i = \langle v, e_i \rangle$, 则

$$\sum_{i=1}^k |\alpha_i|^2 \leq \|v\|^2$$

而且 $v - \sum_{i=1}^k \alpha_i e_i \perp e_i, i = 1, \dots, k$ 。

15. 设 e_1, \dots, e_n 是 n 维内积空间 V 的一组向量, 证明下面的条件等价:

- $\{e_1, \dots, e_n\}$ 是 V 的一组标准正交基;
- 对任意的 $\alpha, \beta \in V$,

$$\langle \alpha, \beta \rangle = \sum_{i=1}^n \langle \alpha, e_i \rangle \langle e_i, \beta \rangle$$

- 对任意的 $\alpha \in V$,

$$\|\alpha\|^2 = \sum_{i=1}^n |\langle \alpha, e_i \rangle|^2$$