



计算方法

刘景铖

计算机软件新技术国家重点实验室
南京大学

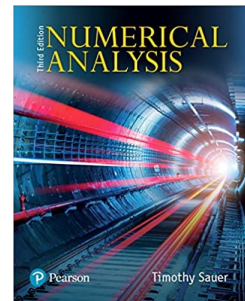


课程基本信息

- 教师：刘景铎
- Email: liu@nju.edu.cn
- Office hour: 周四 14:00-16:00, 计算机系516
- 课程主页: <https://tcs.nju.edu.cn/wiki>
- 数值分析 (Numerical Analysis) (原书第2版)
Timothy Sauer, 机械工业出版社

参考:

- [Numerical Algorithms: Methods for Computer Vision, Machine Learning, and Graphics. Justin Solomon. CRC Press](#)
- $Lx=b$,拉普拉斯方程求解以及它的应用 ($Lx=b$, Laplacian Solver and Their Algorithmic Applications) Nisheeth K. Vishnoi





课程QQ群

- 作业每两周发布一次
- 第二次作业已经发布
- 作业Email: cm_nju_2023@163.com
- QQ群可以讨论课程相关的问题
- 作业可以小组讨论 (≤ 3 人)
 - 清晰标明合作者, 及各自的贡献
 - 使用了的参考资料必须注明
 - 但写作必须由自己独立完成, 不可照抄
 - 类似地, 别人提供的想法也必须注明
- 分辨合作与作弊之间的区别
 - 参与讨论之前, 请先花时间自己进行思考
 - 避免参与你不能提供贡献的讨论
 - 作弊不仅让你丧失一次学习的机会
 - 还会影响你以后的自信心
- 对学术不诚信的行为零容忍

对作弊的认定参照http://www.acm.org/publications/policies/plagiarism_policy



计算方法2023



点击卡片更换背景



保存



分享

群号: 254251456

该群暂不能被搜索, 可前往修改设置



回顾

- 插值：给定数据点，找出多项式经过所有点
 - 对于 n 个数据点，有且仅有一个
 - 存在性：拉格朗日插值方法
 - 唯一性：代数基本定理
 - 龙格现象：等距采样
 - 插值的不可预测性的应用：分享秘密、自纠错码（Reed-Solomon code）
- 函数逼近/拟合：给定函数，找出与之相近的函数（多项式）
 - Weierstrass定理：对于连续函数，多项式可以任意地逼近（即使是等距采样）

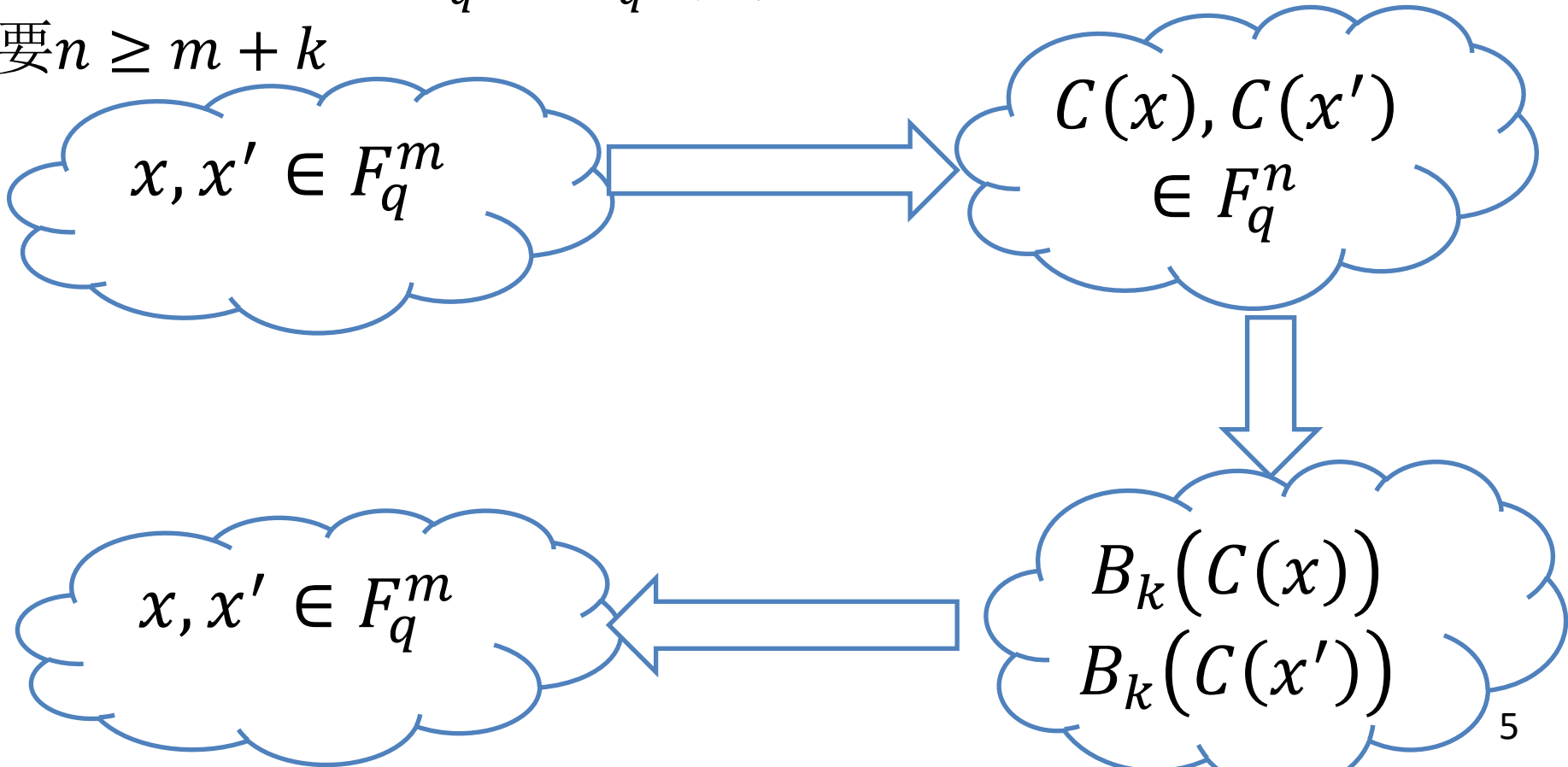
这节课：

- Chebyshev多项式
- 最小二乘法



补充：自纠错码需要多大？ (Hamming bound)

考虑自纠错码 $C: F_q^m \rightarrow F_q^n$. 要处理 k 个字符的 erasure, 那么需要 $n \geq m + k$





补充：自纠错码需要多大？(Hamming bound)

考虑自纠错码 $C: F_q^m \rightarrow F_q^n$. 要处理 k 个字符的 erasure, 那么需要 $n \geq m + k$

反证法：假设 $n \leq m + k - 1$

- 考虑所有的codeword的集合: $\{C(x): x \in F_q^m\}$
- 去掉这些codeword的后面 k 位 (投影到前面 $m - 1$ 位)
- 剩下每个codeword的长度最多为 $m - 1$
- 由Pigeon-hole principle, 一定存在两个codeword在前 $m - 1$ 位是相同的, 记为 x, x'
- 这意味着, 存在 $x, x' \in F_q^m$, $x \neq x'$, 使得 $C(x)$ 和 $C(x')$ 在去掉了最后的 k 位后相等
- 这与能处理 k 个字符的 erasure 的假设矛盾。



补充：自纠错码需要多大？ (Hamming bound)

考虑自纠错码 $C: F_q^m \rightarrow F_q^n$. 要处理 k 个字符的 corruption, 那么需要 $n \geq m + 2k$

反证法：假设 $n \leq m + 2k - 1$

- 考虑所有的codeword的集合: $\{C(x): x \in F_q^m\}$
- 去掉这些codeword的后面 $2k$ 位 (投影到前面 $m - 1$ 位)
- 剩下每个codeword的长度最多为 $m - 1$
- 由Pigeon-hole principle, 一定存在两个codeword在前 $m - 1$ 位是相同的, 记为 x, x'
- 这意味着, 存在 $x, x' \in F_q^m, x \neq x'$, 使得 $C(x)$ 和 $C(x')$ 在去掉了最后的 $2k$ 位后相等
- 因此, 可以从 $C(x)$ 修改最多 k 个字符, 再从 $C(x')$ 修改最多 k 个字符, 将得到一样的信息
- 从接收者的角度看, 将无法区分这是由 $C(x)$ 修改最多 k 个字符得到的, 还是从 $C(x')$ 修改最多 k 个字符得到的
- 这与能处理 k 个字符的 corruption 矛盾

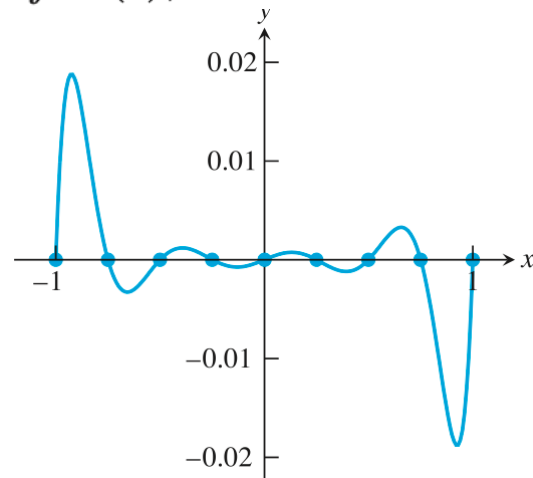


回顾插值中的龙格现象

对连续 n 次可导函数 f ，在 x_1, \dots, x_n 上拉格朗日插值的误差

$$f(x) - P(x) = \frac{(x - x_1)(x - x_2) \cdots (x - x_n)}{n!} f^{(n)}(c),$$

- 证明：令 $q(t)$ 为 x_1, \dots, x_n, x 上插值的多项式
- 则 $q(t) = P(t) + \lambda \prod_{j=1}^n (t - x_j)$ ，其中 $\lambda = \frac{f(x) - P(x)}{\prod_{j=1}^n (x - x_j)}$
- 考虑 $\phi(t) := f(t) - q(t)$ ，则 $\phi(t)$ 在 $n + 1$ 个点处为零
- 由Rolle定理， $\phi'(t)$ 在 n 个点处为零
- 反复应用Rolle定理， $\phi^{(n)}(t)$ 在区间上某个点为零
- 在该点 $\phi^{(n)}(c) = 0 \Rightarrow f^{(n)}(c) = q^{(n)}(c) = \lambda n!$
- 整理即得误差公式



使用等距采样插值时的误差函数

- 如何选点 x_1, \dots, x_n ，才能最小化误差？插值的表现可否媲美函数逼近？
- 采用更多靠近边界的点



Chebyshev插值基点

不失一般性地, 假设函数是定义在 $[-1,1]$ 上的

目标: 选取 x_1, \dots, x_n 使得

$$\max_{x \in [-1,1]} |(x - x_1)(x - x_2) \dots (x - x_n)| \quad (*)$$

最小

- 定理: 令 $x_i = \cos \frac{(2i-1)\pi}{2n}, i = 1, \dots, n$, 得到(*)最小值 $1/2^{n-1}$
- Chebyshev多项式 $T_n(x) := 2^{n-1}(x - x_1)(x - x_2) \dots (x - x_n)$
- Chebyshev插值多项式: 选取Chebyshev基点进行拉格朗日插值得到的多项式



Chebyshev多项式

等价的定义 $T_n(x) := \cos(n \arccos x)$, $x \in [-1, 1]$

这为什么是一个多项式?

- 回忆三角函数的倍角公式
- $\cos(n\theta)$ 可以展开成为关于 $\cos \theta$ 的 n 次多项式
- 令 $x = \cos \theta$, 则 $\cos(n\theta)$ 是关于 x 的 n 次多项式
- $\cos(n\theta) = \cos(n \arccos x)$

$$T_0(x) = 1$$

$$T_1(x) = x$$

$$T_2(x) = \cos 2\theta = 2 \cos^2 \theta - 1 = 2x^2 - 1$$



Chebyshev多项式

等价的定义 $T_n(x) := \cos(n \arccos x), \quad x \in [-1, 1]$

令 $x = \cos \theta$

$$T_0(x) = 1$$

$$T_1(x) = x$$

$$T_2(x) = \cos 2\theta = 2 \cos^2 \theta - 1 = 2x^2 - 1$$

应用倍角公式:

$$T_{n+1}(x) = \cos(n\theta + \theta) = \cos n\theta \cos \theta - \sin n\theta \sin \theta$$

$$T_{n-1}(x) = \cos(n\theta - \theta) = \cos n\theta \cos \theta + \sin n\theta \sin \theta$$

两式相加

$$T_{n+1}(x) + T_{n-1}(x) = 2 \cos n\theta \cos \theta = 2xT_n(x)$$

递归关系

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$



Chebyshev多项式

等价的定义 $T_n(x) := \cos(n \arccos x)$, $x \in [-1, 1]$

一些观察:

- 次数为 n , 最高次项的系数是 2^{n-1}
- $T_n(1) = 1, T_n(-1) = (-1)^n$
- $|T_n(x)| \leq 1$
- $T_n(x)$ 的零点恰好是 $x_i = \cos \frac{(2i-1)\pi}{2n}, i = 1, \dots, n$
 - $\cos n\theta = 0$ iff $n\theta = (2i-1) \cdot \frac{\pi}{2}, i = 1, \dots, n$
- $T_n(x)$ 在-1到1之间来回往返 $n+1$ 次, 分别在 $\cos 0, \cos \frac{\pi}{n}, \dots, \cos \frac{(n-1)\pi}{n}, \cos \pi$



Chebyshev多项式

观察: $T_n(x)$ 在-1到1之间来回往返 $n+1$ 次, 分别在

$$\cos 0, \cos \frac{\pi}{n}, \dots, \cos \frac{(n-1)\pi}{n}, \cos \pi$$

Chebyshev定理: 令 $x_i = \cos \frac{(2i-1)\pi}{2n}, i = 1, \dots, n$, 得到 $\max_{x \in [-1,1]} (x - x_1)(x - x_2) \dots (x - x_n)$ 最小值 $1/2^{n-1}$

证明:

- 假设有另一组 x_i 的选择, 使得 $P_n(x) := (x - x_1)(x - x_2) \dots (x - x_n)$ 在 $[-1,1]$ 上有更小的极大值。换言之, $\forall x \in [-1,1], |P_n(x)| < 1/2^{n-1}$ 。
- 考虑 $P_n(x) - T_n(x)/2^{n-1}$, 则在 $n+1$ 个点上的符号是正负交替的。因此 $P_n(x) - T_n(x)/2^{n-1}$ 有 n 个零点。
- 注意到 $P_n(x)$ 和 $T_n(x)/2^{n-1}$ 都是首1的多项式 (最高次系数为1)
- $P_n(x) - T_n(x)/2^{n-1}$ 的次数最多为 $n-1$ 次, 这与 $P_n(x) - T_n(x)/2^{n-1}$ 有 n 个零点矛盾。(除非 $P_n(x) - T_n(x)/2^{n-1}$ 为零多项式)



函数逼近与插值

Weierstrass 定理: 对于连续函数 f , 多项式可以任意地逼近, 衡量误差的是无穷范数:

$$\|f - p\|_{\infty} := \sup_x |f(x) - p(x)|$$

一般来说, 要找最佳一致逼近的多项式是比较困难的

- 但可以找到与最佳一致逼近的多项式表现相差不大的
- 通过Chebyshev点插值得到的 n 次多项式, 与最佳一致逼近的 n 次多项式之间, 在无穷范数上只差 $O(\log n)$



函数逼近理论在计算机科学

更快的算法:

- 更快地模拟随机游走
- 迭代求解线性方程组 Conjugate gradient; Chebyshev iteration
- 计算最大特征值/特征方向 Lanczos iteration
-

复杂性理论:

- 神经网络(perceptron)表示PARITY的复杂性
- 电路复杂性 circuit complexity
- 通信复杂性 communication complexity
-

课外拓展: 更多的例子可以参见

[The Polynomial Method. In Quantum and Classical Computing](#)

[Faster Algorithms Via Approximation Theory](#)

[Chebyshev Polynomials and Approximation Theory in Theoretical Computer Science and Algorithm Design](#)



函数空间上的线性代数

向量空间: 支持向量加法, 数的乘法, 闭合的线性空间

线性相关: 一组向量是线性相关的, 如果存在它们的一组非零的线性组合为零

例子:

- 实数域上的 n 维向量 R^n
- 实系数的多项式 $R[x]$
 - 加法: 两个多项式相加
 - 数的乘法: 一个常数乘上一个多项式
 - 线性: $c(p(x) + q(x)) = c p(x) + c q(x)$, $(a + b)p(x) = a p(x) + b p(x)$
 - 它的一组基为 $1, x, x^2, x^3, \dots$
 - Chebyshev多项式亦可作为它的一组基, 生成相同的向量空间
- 类似地, 对连续函数, 可定义 $C[a, b]$
- 对 p 阶连续可导函数, 记为 $C^p[a, b]$
- 所有连续函数这个空间非常的大, 考虑一组基生成的函数; 由Weierstrass, 多项式可以任意近似有限区间上的连续函数。
- 称函数族 $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$ 是线性相关的, 如果存在 c_1, c_2, \dots, c_n , 非全为零, 使得
$$c_1 \phi_1(x) + c_2 \phi_2(x) + \dots + c_n \phi_n(x) = 0, \forall x$$

否则称它们是线性无关的



向量空间的范数

R^n 上的无穷范数, 对于向量 $x = (x_1, x_2, \dots, x_n)^T$

$$\|x\|_\infty := \max_i |x_i|$$

R^n 上的2-范数, 对于向量 $x = (x_1, x_2, \dots, x_n)^T$

$$\|x\|_2 := \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}$$

函数 f 的无穷范数 (∞ -norm):

$$\|f\|_\infty := \sup |f|$$

函数 f 的2-范数 (2-norm):

$$\|f\|_2 := \sqrt{\int |f(x)|^2 dx}$$

如果要找到函数 f 的最佳2-范数近似 (最佳平方逼近), 要比最佳无穷范数的近似容易得多: 最小二乘法



最小二乘法—求解线性方程

给定矩阵 $A \in \mathbb{R}^{m \times n}$ ，向量 $b \in \mathbb{R}^m$ ，求 $x \in \mathbb{R}^n$ 使得 $Ax = b$

只有三种情况

- 对任意的向量 b ，存在唯一解
- 不可解，或方程不一致 (inconsistent, over-determined)
- 存在无穷多组解 (under-determined)

性质：只要存在两个不同的解，则一定存在无穷多组解。

接下来：对于方程不一致的情况，我们讨论最小二乘法



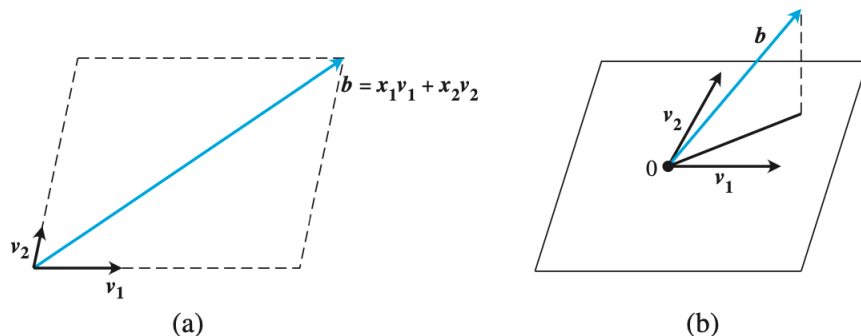
最小二乘法-离散线性版本

从最简单的设定开始。我们想要找出一个线性拟合，使误差的2-范数最小化：

- 给定矩阵 A ，向量 b ，求 x 使得 $\|Ax - b\|_2$ 最小

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ A_1 & A_2 & A_3 & \dots & A_n \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 A_1 + x_2 A_2 + \dots + x_n A_n$$

- 右边是 A_1, A_2, \dots, A_n 能够生成的线性子空间

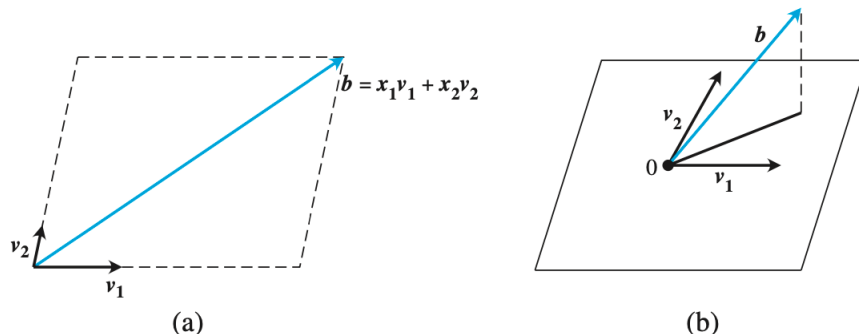


向量 b 不一定在该子空间内，此时，应该找“最接近”的



最小二乘法-几何解释

从最简单的设定开始。我们想要找出一个线性拟合，使误差的2-范数最小化：



最接近的点记为 $A\bar{x}$ ，则 $(A\bar{x} - b) \perp \{A x : x \in R^n\}$

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

要 $\forall x \in R^n$ 同时成立，只有一种可能: $A^T (A\bar{x} - b) = 0$

因此 $A^T A\bar{x} = A^T b$ 法线方程



最小二乘法-微积分解释

注意到最小化 $\|x\|_2$ 和最小化 $\|x\|_2^2$ 是等价的（单调性）

记 $\bar{x} = \operatorname{argmin}_x \|Ax - b\|_2^2$

对 x 求梯度

$$\begin{aligned}\nabla \|Ax - b\|_2^2 &= \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right) \|Ax - b\|_2^2 \\ &= \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right) (x^T A^T A x + b^T b - b^T A x - x^T A^T b) \\ &= 2A^T A x - 2A^T b\end{aligned}$$

把梯度设为0，亦可得到法线方程 $A^T A \bar{x} = A^T b$

注：一般不会直接通过 $\bar{x} = (A^T A)^{-1} A^T b$ 来运算
解线性方程组，与矩阵求逆是两回事，以后会展开讨论



最小二乘法-Gram-Schmidt正交化

- 如果A的每一列是正规化的(orthonormal)?

$$A = \begin{pmatrix} \vdots & \vdots & \vdots & \dots & \vdots \\ A_1 & A_2 & A_3 & \dots & A_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{pmatrix}$$

$$A_i^T A_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

则 $A^T A = I$

法线方程 $A^T A \bar{x} = A^T b$ 简化为 $\bar{x} = A^T b$



最小二乘法-Gram-Schmidt正交化

- 给定A, 如何找出一组正规化的(orthonormal)基?

$$A = \begin{pmatrix} \vdots & \vdots & \vdots & \dots & \vdots \\ A_1 & A_2 & A_3 & \dots & A_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{pmatrix}$$

- $y_1 = A_1, q_1 = y_1/\|y_1\|_2$
- $y_2 = A_2 - q_1(A_2^T q_1), q_2 = y_2/\|y_2\|_2$
- 类似地
- $y_j = A_j - q_1(A_j^T q_1) - q_2(A_j^T q_2) - \dots, q_j = y_j/\|y_j\|_2$



最小二乘法-Gram-Schmidt正交化

- 给定A, 如何找出一组正规化的(orthonormal)基?

$$A = \begin{pmatrix} \vdots & \vdots & \vdots & \dots & \vdots \\ A_1 & A_2 & A_3 & \dots & A_n \\ \vdots & \vdots & \vdots & \dots & \vdots \end{pmatrix}$$

Classical Gram-Schmidt orthogonalization

Let $A_j, j = 1, \dots, n$ be linearly independent vectors.

for $j = 1, 2, \dots, n$

$y = A_j$

for $i = 1, 2, \dots, j - 1$

$r_{ij} = q_i^T A_j$

$y = y - r_{ij} q_i$

end

$r_{jj} = \|y\|_2$

$q_j = y / r_{jj}$

end



最小二乘法-Gram-Schmidt正交化

- 给定A, 如何找出一组正规化的(orthonormal)基?

$$A = \begin{pmatrix} \vdots & \vdots & \vdots & & \vdots \\ A_1 & A_2 & A_3 & \dots & A_n \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix}$$

Modified Gram-Schmidt orthogonalization

Let $A_j, j = 1, \dots, n$ be linearly independent vectors.

```
for  $j = 1, 2, \dots, n$   
   $y = A_j$   
  for  $i = 1, 2, \dots, j - 1$   
     $r_{ij} = q_i^T y$   
     $y = y - r_{ij} q_i$   
  end  
   $r_{jj} = \|y\|_2$   
   $q_j = y / r_{jj}$   
end
```



最小二乘法-方程不一致的情形

- $(A^T A)^{-1} A^T$ 是 A 的一个 *pseudoinverse*

只要 A 的列是线性无关的, 则 $A^T A$ 可逆

$$\bar{x} = (A^T A)^{-1} A^T b$$

思考: 如果是方程数不足呢 (under-determined)?

- $Ax = b$ 的解并不唯一, 有无穷多组解
- 能否找到长度最短的解?
- 下次作业: A 的行是线性无关的话, 亦可写出一个 *pseudoinverse*
- Bonus: 验证这样找到的, 是长度最短的解



内积空间

内积 $\langle \cdot, \cdot \rangle$

- 对称性 $\langle x, y \rangle = \langle y, x \rangle$
- 线性 $\langle a x + b y, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$
- 正定性 $\langle x, x \rangle > 0, \forall x \neq 0$

函数的内积:

$$\langle f, g \rangle := \int f(x)g(x)dx$$

带权重的函数内积

$$\langle f, g \rangle_w := \int f(x)g(x)w(x)dx$$

内积天然自带范数的定义: $\|x\| = \sqrt{\langle x, x \rangle}$



正交函数族与正交多项式

称函数族 $\phi_1(x), \phi_2(x), \dots$ 是正交的, 如果它们满足:

$$\langle \phi_i, \phi_j \rangle = \begin{cases} C_i, & i = j \\ 0, & i \neq j \end{cases}$$

- 如果 $\phi_1(x), \phi_2(x), \dots$ 是多项式, 也称之为正交多项式

Chebyshev多项式也是多项式的一组基

它关于 $w(x) = \frac{1}{\sqrt{1-x^2}}$ 所定义的内积是正交的:

$$\langle f, g \rangle_w := \int_{-1}^1 f(x)g(x) \frac{dx}{\sqrt{1-x^2}}$$



下节课

傅里叶变换
三角插值