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

回顾

上节课：

- Chebyshev插值
- Chebyshev多项式
- 由函数/多项式亦可组成的向量空间

这节课：

- **最小二乘法**-离散线性版本：给定矩阵A，向量b，求x使得$\|Ax-b\|_2^2$**最小**
- FFT

课程计划

最小二乘法

- **建模和应用讨论：图像拼接**
- **三角级数近似**

傅里叶变换 (Fourier Transform)

- 复数、单位根
- 离散傅里叶变换 (Discrete Fourier Transform, DFT)
- 快速傅里叶变换 (Fast Fourier Transform, FFT)
- 复杂性分析
- 快速傅里叶变换的逆变换(Inverse FFT)

最小二乘法建模：图像拼接

照片来源：由本人拍摄于2022-03-18

最小二乘法建模：图像拼接

- 在图片上选取一些参考点（比如同一个建筑，对应的树）
- 希望找到矩阵$A \in \mathbb{R}^{2\times 2}$和向量 $b \in \mathbb{R}^2$
- 使得对于每一个参考点都有

$$\approx A \times$$

$$+ b$$

- **未知量是什么？**
- **思考：值得注意的是这里的量词的顺序**
  - 如果，对于每一个参考点，都要求找到一个矩阵$A \in \mathbb{R}^{2\times 2}$和向量 $b \in \mathbb{R}^2$**最小化误差**呢？

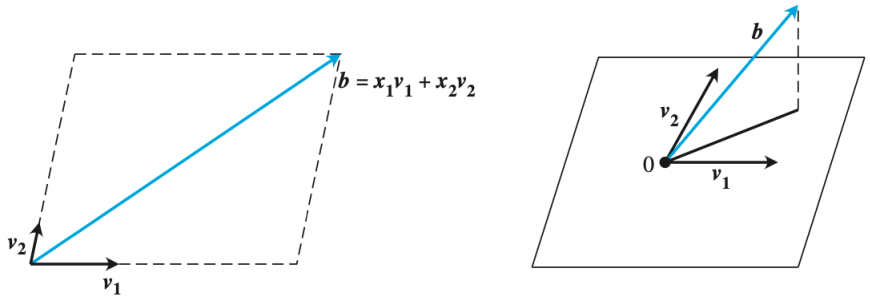
最小二乘法-离散线性版本

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$\|Ax-b\|_2^2$最小

$$\begin{pmatrix} \vdots & \vdots & \vdots & & \vdots \\ A_1 & A_2 & A_3 & \dots & A_n \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1A_1 + x_2A_2 + \cdots + x_nA_n$$

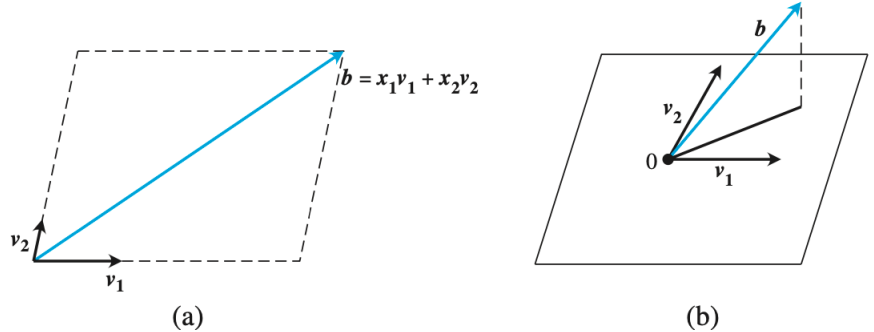
- 右边是 $A_1, A_2, \dots, A_n$能够生成的线性子空间



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

对x求梯度

$$\nabla\|Ax - b\|_2^2 = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}\right)||Ax - b||_2^2$$

$$= \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}\right)(x^T A^T A x + b^T b - b^T A x - x^T A^T b)$$

$$= 2A^T A\, x - 2A^T b$$

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-变分法解释

记误差函数为$E(x) := \|Ax - b\|_2^2$

则$\bar{x}$满足法线方程$A^T A\bar{x} = A^T b$当且仅当

$$\forall y, E(\bar{x}) \leq E(\bar{x} + y)$$

证明(Sketch)：

$$E(\bar{x} + y) = E(x) + (Ay)^\top (Ay) + 2y^\top A^\top (A\bar{x} - b)$$

注意到

$$(\forall y, y^\top A^\top (A\bar{x} - b) = 0) \iff A^\top (A\bar{x} - b) = 0$$

考虑逆反命题：

$$(\exists y, y^\top A^\top (A\bar{x} - b) \neq 0) \Leftrightarrow \left(\exists y, y^\top A^\top (A\bar{x} - b) < 0\right)$$

进而，通过取足够短的向量y，可以令

$$(Ay)^\top (Ay) + 2y^\top A^\top (A\bar{x} - b) < 0$$

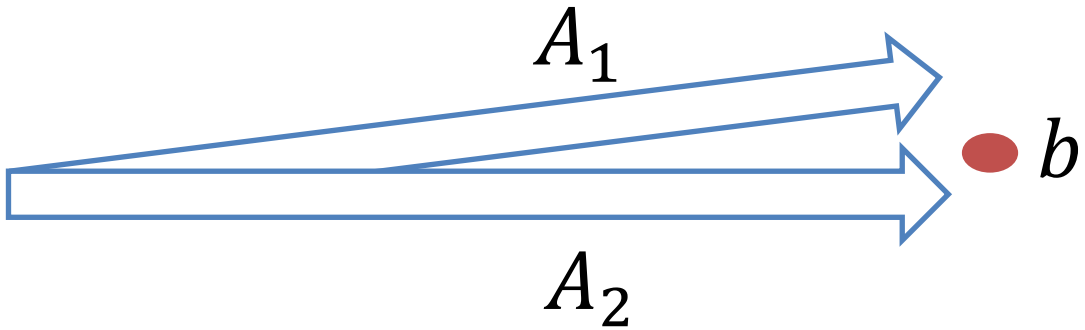
进而得到$E(\bar{x}) > E(\bar{x} + y)$

最小二乘法与正交性

给定矩阵A，向量b，求x使得$\|Ax-b\|_2^2$最小。

由法线方程 $A^TA\bar{x}=A^Tb$，得出 $\bar{x}=(A^TA)^{-1}A^Tb$

- 但是，一般数值算法中不会直接求解法线方程
- 更不会求$(A^TA)^{-1}$



- 如图所示，不管选A_1还是 A_2，都能给出对b不错的近似，但是对应的x是非常不一样的

最小二乘法-Gram-Schmidt正交化

一个更好的做法是通过Gram-Schmidt求$A = QR$分解

- 如果A的每一列是正规化的(orthonormal)？

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

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

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？

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

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

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？

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

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

最小二乘法与正交性

通过Gram-Schmidt求$A = QR$分解

$$\begin{pmatrix} \vdots & \vdots & \vdots & & \vdots \\ A_1 & A_2 & A_3 & \cdots & A_n \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \vdots & & \vdots \\ q_1 & q_2 & q_3 & \cdots & q_n \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & r_{nn} \end{pmatrix}$$

- 其中Q为正交阵：$Q^T Q = I$
  - 这意味着$\langle q_i, q_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$
- R为上三角阵

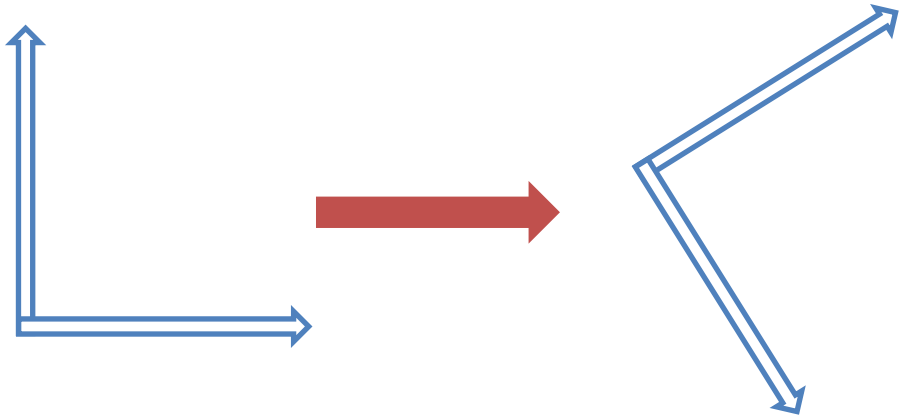
通过QR分解， $A^T A\bar{x} = A^T b \Rightarrow \bar{x} = R^{-1} Q^T b$

并不需要计算 $A^T A$

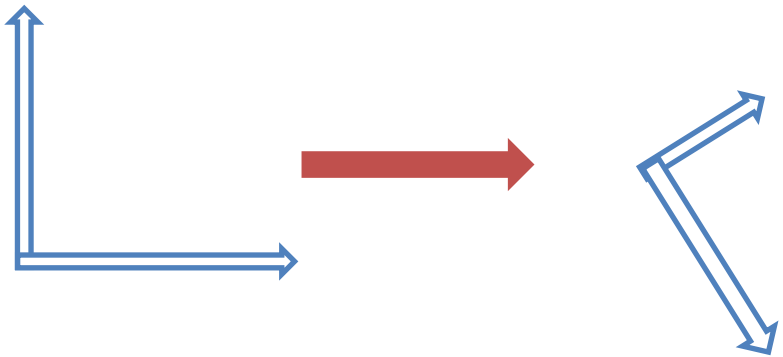
实践中，更常用，更数值稳定的是Householder反射子，而不是Gram-Schmidt

最小二乘法与正交性

正交变换



非正交变换



最小二乘法与正交性

若Q为正交阵：$Q^TQ = I$，则$\|Qx\|_2 = \|x\|_2$

Isometry：保距映射

思考：除了长度，还可以考虑角度

$$\langle Qx, Qy \rangle = ?$$

最小二乘法-方程不一致的情形

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

只要A的列是线性无关的，则A^TA可逆

$$\bar{x} = (A^TA)^{-1}A^Tb$$

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$\phi_1(x), \phi_2(x), \ldots$是正交化的，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 $\phi_1(x), \phi_2(x), \ldots$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：

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

正交系统中的最小二乘法：傅里叶级数

给定函数 f 和一组基函数 ϕ_i，能否找到系数 $\{c_i\}$ 使得

$$f(x) \approx \sum_i c_i\, \phi_i(x)$$

设 $\phi_0 = \frac{1}{2}$,

$$\phi_k(x) = \cos(kx)\,, k = 1,2,\dots,n,$$
$$\phi_{n+k}(x) = \sin(kx)\,, k = 1,2,\dots,n-1$$

则函数族 $\phi_0(x), \phi_1(x), \dots$ 在 $[-\pi,\pi]$ 上是正交的

证明： $\cos(kx)$ 和 $\sin(kx)$ 在 $[-\pi,\pi]$ 上的积分都为0，再由积化和差公式即可得：

$$\sin(y)\cos(x) = \frac{1}{2}(\sin(y-x) + \sin(y+x))$$
$$\cos(y)\cos(x) = \frac{1}{2}(\cos(y-x) + \cos(y+x))$$
$$\sin(y)\sin(x) = \frac{1}{2}(\cos(y-x) - \cos(y+x))$$

正交系统中的最小二乘法：傅里叶级数（选讲）

给定函数 f，能否找到系数$\{a_i, b_i\}$使得

$$f(x) \approx \frac{a_0}{2} + a_n \cos nx + \sum_{k=1}^{n-1} (a_k \cos kx + b_k \sin kx)$$

这里也可以使用最小二乘法（sketch）：

- 记误差函数为

$$E(x) := f(x) - \left(\frac{a_0}{2} + a_n \cos nx + \sum_{k=1}^{n-1} (a_k \cos kx + b_k \sin kx) \right)$$

- 考虑 $\|E(x)\|_2^2 = \langle E(x), E(x) \rangle = \int_{-\pi}^{\pi} E(x)^2 dx$
- 对 $\|E(x)\|_2^2$关于系数 $\{a_i, b_i\}$求偏导并设为0，即得到法线方程
- 由于正交性，可推导得

$$a_k = \frac{\langle f, \cos kx \rangle}{\langle \cos kx, \cos kx \rangle}$$

$$b_k = \frac{\langle f, \sin kx \rangle}{\langle \sin kx, \sin kx \rangle}$$