线性代数 (Linear Algebra)

第六章 Orthogonality and Least Squares

§ 6.4 The Gram-Schmidt Process

Gram-Schmidt 过程

衡益

2020 年 12 月 23 日，中山大学南校区

Gram-Schmidt 过程

Gram-Schmidt 过程

- 什么是**Gram – Schmidt** 过程?

目标：找到给定空间W的一组正交基

前提：已知该空间的一组非正交基

3

Gram-Schmidt 过程

Gram-Schmidt过程是用于产生$\mathbb{R}^n$的任何非零子空间的正交或标准正交基的简单算法。

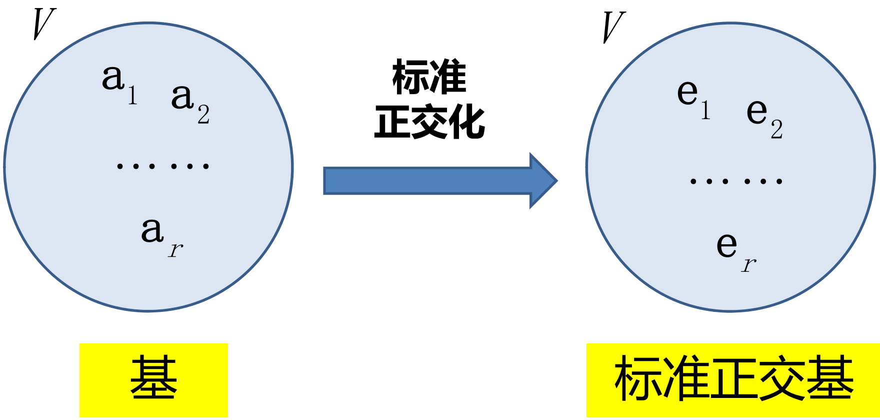
向量形式 → **标准正交化**

矩阵形式 → **QR分解**

4

Gram-Schmidt 过程

向量形式



5

Gram-Schmidt 过程

定理 对$\mathbb{R}^n$中子空间W的一个基$\{x_1, \ldots, x_p\}$，定义：

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

$$\vdots$$

$$v_p = x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}$$

那么$\{v_1, \ldots, v_p\}$是W的一个正交基，此外

$\mathrm{Span}\{v_1, \ldots, v_k\} = \mathrm{Span}\{x_1, \ldots, x_k\}$，其中$1 \le k \le p$

6

Gram-Schmidt 过程

定理 对$1 \le k \le p$，取$W_k = \{x_1, ..., x_k\}$.

证明 令$v_1 = x_1$，则有$\text{Span}\{v_1\} = \text{Span}\{x_1\}$.

若对$k < p$，构造$v_1, ..., v_k$，使得$\{v_1, ..., v_k\}$是W_k的一个正交基，

定义$v_{k+1} = x_{k+1} - \text{proj}_{W_k} x_{k+1}$，

注意向量$\text{proj}_{W_k} x_{k+1}$属于$W_k$，因而属于$W_{k+1}$，

故v_{k+1}也属于W_{k+1}(因为W_{k+1}是子空间，且减法封闭).

更进一步，由于x_{k+1}不属于W_k，可得$v_{k+1} \neq 0$.

因而$\{v_1, ..., v_{k+1}\}$是$(k+1)$维空间W_{k+1}中非零向量形成的正交基.

由4.5节基定理可知，这个集合是W_{k+1}的正交基，

从而$W_{k+1} = \text{Span}\{v_1, ..., v_{k+1}\}$当$k+1 = p$时成立，

归纳证明结束√.

Gram-Schmidt 过程

4.5节基定理回顾

(基定理)

✓ 若V是一个p维向量空间，$p \ge 1$，V中任意含有p个元素的线性无关集必然是V的一个基。任意含有p个元素且生成V的集合必然是V的一个基。

Gram-Schmidt 过程

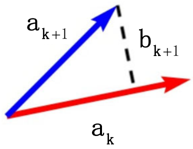
$$a_1 \qquad a_2 \qquad \cdots \qquad a_r$$

$$b_1 = a_1$$

$$b_2 = a_2 - \frac{[b_1, a_2]}{[b_1, b_1]} b_1$$

$$\cdots\cdots$$

$$b_r = a_r - \frac{[b_1, a_r]}{[b_1, b_1]} b_1 - \frac{[b_2, a_r]}{[b_2, b_2]} b_2 - \cdots - \frac{[b_{r-1}, a_r]}{[b_{r-1}, b_{r-1}]} b_{r-1}$$



$$b_{k+1} = a_{k+1} - \mathrm{proj}(a_{k+1})$$

$$[b_1, b_1] = \|b_1\|^2$$

Schmidt orthogonalization

9

Gram-Schmidt 过程

标准基

$$e_1 = \frac{1}{\|b_1\|} b_1$$

$$e_2 = \frac{1}{\|b_2\|} b_2$$

$$\vdots$$

$$e_r = \frac{1}{\|b_r\|} b_r$$

$\cdots$

a_1
a_2
$\vdots$
a_r

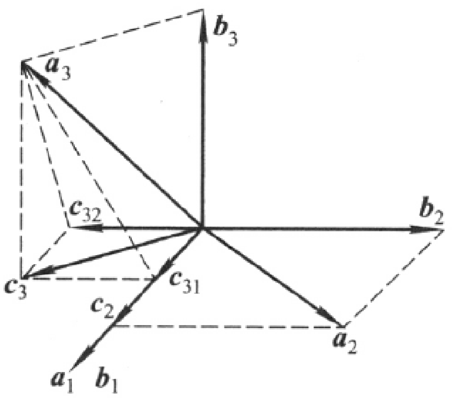
10

Gram-Schmidt 过程举例

设 $a_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, a_2 = \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}, a_3 = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix}$，用Gram-Schmidt过程将该向量组正交化

解：

$b_1 = a_1 =$



$= \frac{5}{3}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$

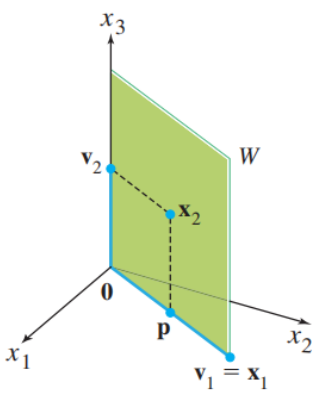
$b_3 = a_3 - [\varepsilon$

$\begin{matrix} -1 \\ 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

$\Rightarrow e_1 = \frac{1}{\sqrt{6}}$

Gram-Schmidt 过程举例

设 $W = \mathrm{span}\{x_1, x_2\}$，$x_1 = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$，$x_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$，利用Schmidt为$W$构造正交基$\{v_1, v_2\}$.



令$v_1 = x_1$，

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1}{\mathbf{x}_1 \cdot \mathbf{x}_1}\mathbf{x}_1$$

$$= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Gram-Schmidt 过程举例

MATLAB算法-向量形式

```
function v=Schmidt_orthogonalization(a)
[m,n] = size(a);
b=zeros(m,n);
%正交化
b(:,1)=a(:,1);
for i=2:n
   for j=1:i-1
      b(:,i)=b(:,i)-dot(a(:,i),b(:,j))/dot(b(:,j),b(:,j))*b(:,j);
   end
   b(:,i)=b(:,i)+a(:,i);
end
%单位化
for k=1:n
     v(:,k)=b(:,k)/norm(b(:,k));
end
```


Gram-Schmidt 过程

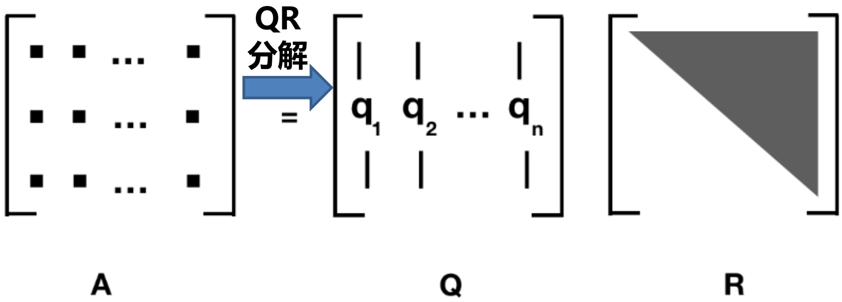
定理

如果$m \times n$矩阵A的列线性无关，那么A可以分解为A=QR，其中Q是一个$m \times n$矩阵，其列形成Col A的一个标准正交基，R是一个$m \times n$上三角可逆矩阵且在对角线上的元素为正数.

可以看做是Gram-schmidt过程的推论

Gram-Schmidt 过程

矩阵形式　　**Q：标准正交矩阵; R：上三角矩阵**



$$Q = [\mathbf{e}_1, \ldots, \mathbf{e}_n] \qquad R = \begin{pmatrix} \langle \mathbf{e}_1, \mathbf{a}_1 \rangle & \langle \mathbf{e}_1, \mathbf{a}_2 \rangle & \langle \mathbf{e}_1, \mathbf{a}_3 \rangle & \ldots \\ 0 & \langle \mathbf{e}_2, \mathbf{a}_2 \rangle & \langle \mathbf{e}_2, \mathbf{a}_3 \rangle & \ldots \\ 0 & 0 & \langle \mathbf{e}_3, \mathbf{a}_3 \rangle & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Q的计算参考后两页

Gram-Schmidt 过程举例

例题

$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$，利用Gram-Schmidt正交化对A进行QR分解

Gram-Schmidt 过程举例

解析

利用Gram-Schmidt正交化求得A的标准向量基，

$$\therefore\ Q=\begin{pmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{pmatrix},$$

$$\because\ Q^TA = Q^T(QR) = IR = R,$$

$$\therefore\ R = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{pmatrix}.$$

Gram-Schmidt 过程举例

MATLAB算法-矩阵形式

```
for j=1:n                    % Gram-Schmidt orthogonalization
  v=A(:,j);                  % v begins as column j of A
  for i=1:j-1
    R(i,j)=Q(:,i)'*A(:,j);   % modify A(:,j) to v for more accuracy
    v=v-R(i,j)*Q(:,i);       % subtract the projection (q_i^T a_j)q_i = (q_i^T v)q_i
  end                        % v is now perpendicular to all of q_1,...,q_{j-1}
  R(j,j)=norm(v);
  Q(:,j)=v/R(j,j);           % normalize v to be the next unit vector q_j
end
```

线性代数 (Linear Algebra)

第六章 Orthogonality and Least Squares

§ 6.5 Least Squares Problems

最小二乘问题

衡益

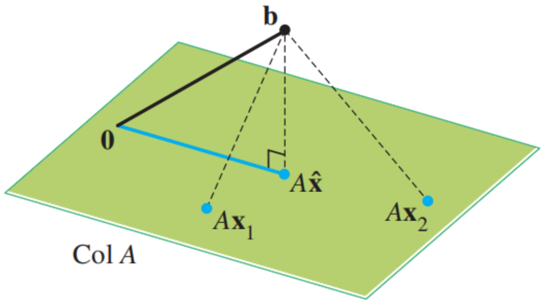
2021 年 12 月 23 日，中山大学南校区

最小二乘估计

最小二乘估计

定义 对于$A_{m\times n}$和$b \in R^m$，方程$Ax = b$的最小二乘解是$\hat{x}$，使得：

$$\|b - A\hat{x}\| \le \|b - Ax\|.$$



最小二乘估计

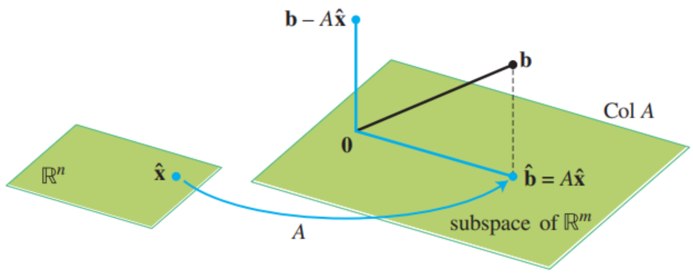
推导 假设$\hat{x}$满足$A\hat{x} = \hat{b}$.

由正交分解定理可知$b-\hat{b}$正交于A的列向量，

即$b-A\hat{x}$正交于A的所有列向量，故

$A^T(b-A\hat{x}) = 0.$

$\Rightarrow A^TA\hat{x} = A^Tb$



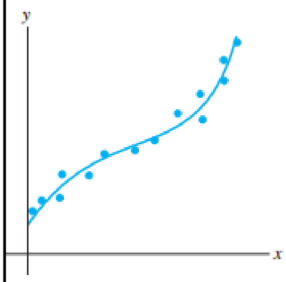
最小二乘估计

定理

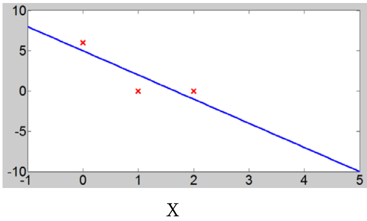
由上页推导可知，

$Ax = b$ 的最小二乘解与 $A^{T}A\hat{x} = A^{T}b$ 的非空解集一致

$$A^{T}A\hat{x} = A^{T}b$$



线性方程



非线性方程

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

对于QR分解：

$$\hat{x} = R^{-1}Q^{T}b$$

最小二乘估计

定理

矩阵$A^{T}A$是可逆的充分必要条件是：

A的列是线性无关的，在这种情形下，

方程$Ax=b$有唯一最小二乘解$\hat{x}$且它有下面的表示：

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

最小二乘估计

线性方程例题

举例1： 找出离点 (0, 6)，(1, 0) 和 (2, 0) 最“近”的线。

设有直线 $b = C + Dt$ 经过这三点

$$\Rightarrow A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, x = \begin{pmatrix} C \\ D \end{pmatrix}, b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix},$$

$Ax = b$ 无解

$$A^T Ax = A^T b \Rightarrow \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$$

$$\Rightarrow C = 5, D = -3$$

25

最小二乘估计

线性方程例题

举例2：

$$A=\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, b = \begin{pmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{pmatrix},$$ 求 $Ax = b$ 的最小二乘解

26

最小二乘估计

线性方程例题

解：

$$A^TA=\begin{pmatrix}6&2&2&2\\2&2&0&0\\2&0&2&0\\2&0&0&2\end{pmatrix},\quad A^Tb=\begin{pmatrix}4\\-4\\2\\6\end{pmatrix}$$

$A^TAx=A^Tb$ 的增广矩阵为：

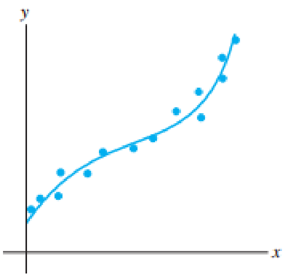
$$\begin{pmatrix}6&2&2&2&4\\2&2&0&0&-4\\2&0&2&0&2\\2&0&0&2&6\end{pmatrix}\sim\begin{pmatrix}1&0&0&1&3\\0&1&0&-1&-5\\0&0&1&-1&-2\\0&0&0&0&0\end{pmatrix}\quad 故：\hat{x}=\begin{pmatrix}3\\-5\\-2\\0\end{pmatrix}+x_4\begin{pmatrix}-1\\1\\1\\1\end{pmatrix},$$

即 $x_1=3-x_4$，$x_2=-5+x_4$，$x_3=-2+x_4$，其中 x_4 是不受约束的.

27

最小二乘估计

非线性方程



目的：我们尝试将数据组

$$(x_1,y_1),\cdots,(x_k,y_k)$$

拟合到某个多项式 $y=a_0+a_1x+a_2x^2+\cdots+a_nx^n$ 上。

理想中，我们希望

$$\begin{aligned}a_0+a_1x_1+\cdots+a_nx_1^n&=y_1\\a_0+a_1x_2+\cdots+a_nx_2^n&=y_2\\&\vdots\\a_0+a_1x_k+\cdots+a_nx_k^n&=y_k\end{aligned}$$

28

最小二乘估计

非线性方程

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

观测向量　　设计矩阵　　系数矩阵

$$\vec{b} = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & x_k^2 & \cdots & v_k^n \end{bmatrix} \qquad \vec{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}$$

29

最小二乘估计

非线性方程

现在只需解方程

$$X\vec{a} = \vec{b}$$

但是基于数据组， $X\vec{a} = \vec{b}$ 可能矛盾，因此我们尝试寻找一个最佳拟合的多项式（也就是当 $||X\vec{a} - \vec{b}||$ 被最小化时的多项式）。

根据之前提到的正则系引理，我们看到 $||X\vec{a} - \vec{b}||$ 被最小化当且仅当

$$X^T X\vec{a} = X\vec{b}$$

30

最小二乘估计

最小二乘与QR分解法

给定一个$m \times n$矩阵A，且具有线性无关的列，取$A = QR$，那么对每一个属于$\mathbb{R}^m$的b，矩阵$Ax = b$有唯一的最小二乘解，其解为：

$\hat{x}=R^{-1}Q^Tb$

取$\hat{x} = R^{-1}Q^Tb$，

那么$A\hat{x} = QR\hat{x} = QRR^{-1}Q^Tb = QQ^Tb$

31

最小二乘估计

最小二乘与QR分解法MATLAB代码

```
[m,n] = size(A);
[Q,R,P] = qr(A);
c = Q'*b;
% Determine rank of A (as before).
tol = max(size(A))*eps*abs(R(1,1));
r = 1;
while ( abs(R(r+1,r+1)) >= tol & r < n ); r = r+1; end
% Solve least squares problem to get y2
S = [ R(1:r,1:r) \ R(1:r,r+1:n);
eye(n-r) ];
t = [ R(1:r,1:r) \ c(1:r);
zeros(n-r,1) ];
y2 = S \ t; % solve least squares problem using backslash
% Compute x
y1 = R(1:r,1:r) \ ( c(1:r) - R(1:r,r+1:n) * y2 );
x = P*[y1;y2];
```

32

最小二乘估计

QR分解例题

$$A=\begin{pmatrix}1&3&5\\1&1&0\\1&1&2\\1&3&3\end{pmatrix},\ b=\begin{pmatrix}3\\5\\7\\-3\end{pmatrix},\quad \text{用QR分解和最小二乘法求}Ax=b$$

33

最小二乘估计

解析

$$\because\ A=QR=\begin{pmatrix}1/2&1/2&1/2\\1/2&-1/2&-1/2\\1/2&-1/2&1/2\\1/2&1/2&-1/2\end{pmatrix}\begin{pmatrix}2&4&5\\0&2&3\\0&0&2\end{pmatrix},$$

$$Q^Tb=\begin{pmatrix}1/2&1/2&1/2&1/2\\1/2&-1/2&-1/2&1/2\\1/2&-1/2&1/2&-1/2\end{pmatrix}\begin{pmatrix}3\\5\\7\\-3\end{pmatrix}=\begin{pmatrix}6\\-6\\4\end{pmatrix},$$

$$\because\ R\hat{x}=Q^Tb,$$

$$\therefore\ \begin{pmatrix}2&4&5\\0&2&3\\0&0&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}6\\-6\\4\end{pmatrix},\ \hat{x}=\begin{pmatrix}10\\-6\\2\end{pmatrix}.$$

34

模型参数估计

数学模型 $y(t) = m_1 + m_2 t - \frac{1}{2} m_3 t^2$

含噪数据

$t = 1,\ y = 106.1$ (105.1)
$t = 2,\ y = 189.4$ (190.4)
$t = 3,\ y = 266.9$ (265.9)

$$\underbrace{\begin{pmatrix} 1 & 1 & -0.5 \\ 1 & 2 & -2.0 \\ 1 & 3 & -4.5 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 106.1 \\ 189.4 \\ 266.9 \end{pmatrix}}_{b}$$

逆矩阵求解 $x = A^{-1}b$

$$\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} 17 \\ 92 \\ 5.8 \end{pmatrix}$$

$m_1 = 10$
$m_2 = 100$
$m_3 = 9.8$

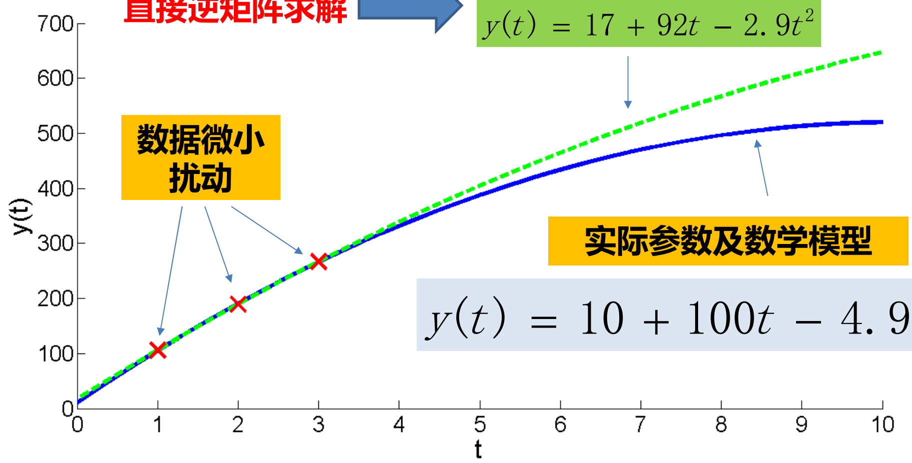
测试参数!

35

模型参数估计

直接逆矩阵求解 → **估算参数及数学模型**

$y(t) = 17 + 92t - 2.9t^2$



实际参数及数学模型

$y(t) = 10 + 100t - 4.9t^2$

36

模型参数估计

含噪数据

$t=1,\ y=106.1;\quad t=4,\ y=332.6;$

$t=2,\ y=189.4;\quad t=5,\ y=387.7;$

$t=3,\ y=266.9;\quad t=6,\ y=432.6;$

$$A=\begin{pmatrix}1 & 1 & -0.5\\ 1 & 2 & -2.0\\ 1 & 3 & -4.5\\ 1 & 4 & -8.0\\ 1 & 5 & -12.5\\ 1 & 6 & -18.0\end{pmatrix},\quad x=\begin{pmatrix}m_1\\ m_2\\ m_3\end{pmatrix},\quad b=\begin{pmatrix}106.1\\ 189.4\\ 266.9\\ 332.6\\ 387.7\\ 432.6\end{pmatrix}$$

最小二乘解 (Least Square)

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

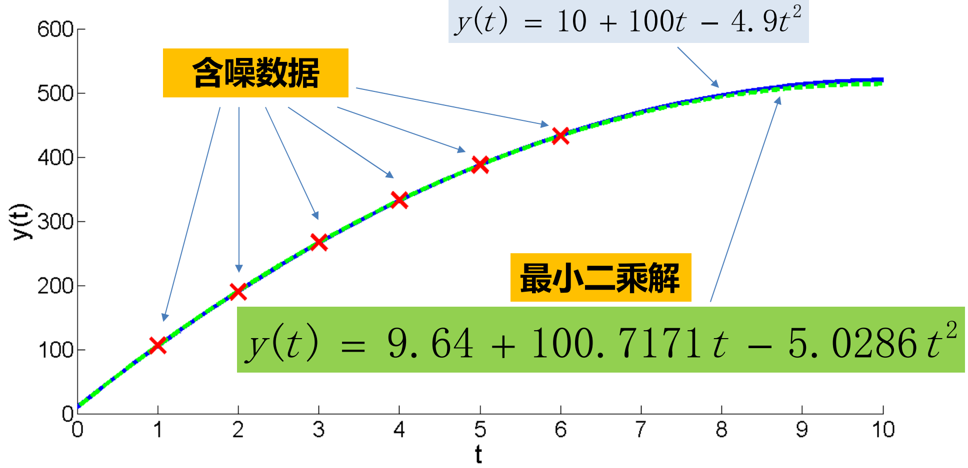
$$\begin{pmatrix}m_1\\ m_2\\ m_3\end{pmatrix}=\begin{pmatrix}9.64\\ 100.7171\\ 10.0571\end{pmatrix}$$

37

模型参数估计

目标参数及数学模型

$y(t)=10+100t-4.9t^2$



$$y(t)=9.64+100.7171\,t-5.0286\,t^2$$

38

内积空间

内积空间

定义

内积空间是具有内积运算的线性空间。

在向量空间V中，对向量u, v内积运算满足下列公理：

1. $\langle u, v\rangle = \langle v, u\rangle$
2. $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
3. $\langle cu, v\rangle = c\langle u, v\rangle$
4. $\langle u, u\rangle \geq 0$，当且仅当u=0时$\langle u, u\rangle = 0$

内积空间举例

例题1. $\langle u, v\rangle = 4u_1v_1 + 5u_2v_2$，证明该式定义的是内积运算

证明:1. $\langle u, v\rangle = 4u_1v_1 + 5u_2v_2 = 5u_2v_2 + 4u_1v_1 = \langle v, u\rangle$

2. $\langle u + v, w\rangle = 4(u_1 + v_1)w_1 + 5(u_2 + v_2)w_2$

$= 4u_1w_1 + 4v_1w_1 + 5u_2w_2 + 5v_2w_2$

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

3. $\langle cu, v\rangle = 4(cu_1)v_1 + 5(cu_2)v_2$

$= c(4u_1v_1 + 5u_2v_2)$

$= c\langle u, v\rangle$

4. $\langle u, u\rangle = 4u_1^2 + 5u_2^2 \geq 0$,

当且仅当$u_1 = u_2 = 0$时$4u_1^2 + 5u_2^2 = 0$.

内积空间

定义

内积空间中：

1. 范数(norm)：$\|v\| = \sqrt{\langle v, v\rangle}$

2. u, v间距离(distance)：$\|u - v\|$

3. $u \perp v$：$\langle u, v\rangle = 0$

内积空间

例题

定义内积空间P_n，对于P_n中的p, q，满足：

$\langle p, q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n)$.

令$p(t) = 12t^2, q(t) = 2t - 1$.

$t_0 = 0, t_1 = \frac{1}{2}, t_2 = 1$, 求向量$p, q$的长度.

解：$\|p\|^2 = \langle p, p\rangle = [p(0)]^2 + [p(\frac{1}{2})]^2 + [p(1)]^2$

$= 0 + [3]^2 + [12]^2 = 153,$

$\therefore \|p\| = \sqrt{153}.$

同理 $\langle q, q\rangle = 2, \|q\| = \sqrt{2}.$

43

内积空间

例题

同上对于内积空间P_2,

令$p(t) = 1, q(t) = t, r(t) = t^2$.

$t_0 = -2, t_1 = -1, t_2 = 0, t_3 = 1, t_4 = 2,$

利用Gram-Schmidt正交化构造P_2空间的正交基.

44

内积空间

解：$\because\ p=\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix},\ q=\begin{bmatrix}-2\\-1\\0\\1\\2\end{bmatrix},\ r=\begin{bmatrix}4\\1\\0\\1\\4\end{bmatrix}.$

构造P_2空间正交基v_0, v_1, v_2，

令$v_0=p$，$\because <p, q>=0, \therefore v_1=q$，

由Gram-Schmidt可得 $v_2=\begin{bmatrix}2\\-1\\-2\\-1\\2\end{bmatrix}.$

45

内积空间性质

性质 Cauchy-Schwarz不等式：$|<u, v>| \le \|u\|\,\|v\|$

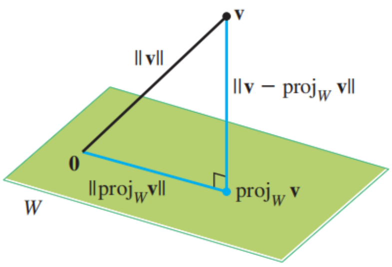
证明：(1)u = 0时，上式显然成立；

(2)u ≠ 0时，设W是u张成的子空间，有：

$$\|\mathrm{proj}_W v\| = \left\|\frac{<v, u>}{<u, u>}u\right\| = \frac{|<v, u>|}{|<u, u>|}\|u\| = \frac{|<v, u>|}{\|u\|^2}\|u\| = \frac{|<u, v>|}{\|u\|}.$$

$$\because \|\mathrm{proj}_W v\| \le \|v\|, \therefore \frac{|<u, v>|}{\|u\|} \le \|v\|,$$

$|<u, v>| \le \|u\|\,\|v\|$ 得证.



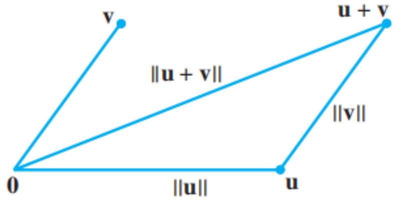
内积空间性质

性质

三角不等式：$\|u+v\| \le \|u\| + \|v\|$

证明：

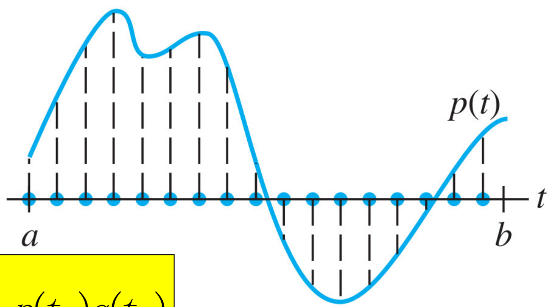
$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\ &\le \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\ &\le \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 \qquad \text{Cauchy–Schwarz} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2\end{aligned}$$



47

内积空间与微积分

内积与微积分



$$< p, q >= \frac{1}{n+1}\sum_{j=0}^{n} p(t_j)q(t_j)$$

$$= \frac{1}{b-a}[\sum_{j=0}^{n} p(t_j)q(t_j)\Delta t]$$

$$\Rightarrow \quad \frac{1}{b-a}\int_a^b p(t)q(t)dt$$

48

内积空间与微积分

例题

定义内积空间$C[0,1]$，对于$f, g \in C[0,1]$满足：

$\langle f, g\rangle=\int_0^1 f(t)g(t)dt, \langle f, f\rangle=\int_0^1 [f(t)]^2 dt \geq 0;$

设W为$C[0,1]$的子空间，由$p_1(t)=1, p_2(t)=2t-1,$

$p_3(t)=12t^2$.利用Grant-Schmidt正交化过程求W空间正交基。

内积空间与微积分

解：令$q_1=p_1$,

$\because \langle p_2, q_1\rangle=\int_0^1 (2t-1)(1)dt=(t^2-t)\Big|_0^1=0,$

$\therefore$ 令$q_2=p_2$.

$\langle p_3, q_1\rangle=\int_0^1 (12t^2)(1)dt=4, \langle q_1, q_1\rangle=\int_0^1 (1)(1)dt=1,$

$\langle p_3, q_2\rangle=\int_0^1 (12t^2)(2t-1)dt=2, \langle q_2, q_2\rangle=\int_0^1 (2t-1)^2 dt=1/3.$

$$\therefore \operatorname{proj}_W p_3=\frac{\langle p_3, q_1\rangle}{\langle q_1, q_1\rangle}q_1+\frac{\langle p_3, q_2\rangle}{\langle q_2, q_2\rangle}q_2=4q_1+6q_2,$$

$\therefore q_3=p_3-4q_1-6q_2=12t^2-12t+2.$

内积空间应用

加权最小二乘法

实际值：y，拟合值：$\hat{y}$；

均方误差(sum of the squanres for error)：

$SS(E)=(y_1-\hat{y}_1)^2+\cdots+(y_n-\hat{y}_n)^2$；

加权均方误差：

$WSS(E)=w_1^2(y_1-\hat{y}_1)^2+\cdots+w_n^2(y_n-\hat{y}_n)^2$；

51

内积空间应用

加权最小二乘法

$$Wy=\begin{pmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & w_n \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}=\begin{pmatrix} w_1y_1 \\ w_2y_2 \\ \vdots \\ w_ny_n \end{pmatrix}$$

$$\Downarrow$$

$$\|Wy-W\hat{y}\|^2=\|Wy-WA\hat{x}\|^2$$

$$\Downarrow$$

$$WAx=Wy$$

$$\Downarrow$$

$$(WA)^TWAx=(WA)^TWy$$

52

内积空间应用

例题

找数据点(–2, 3) , (–1, 5) , (0, 5) , (1, 4) , (2, 3)的最优拟合线y=$\beta_0 + \beta_1 x$.由于最后两个点的测量误差较大，权重设为其他点的一半。

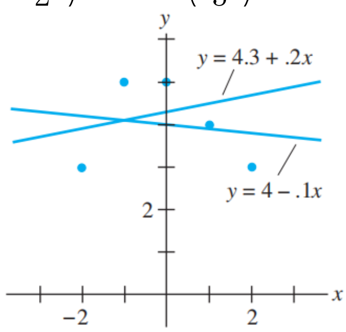
内积空间应用

解：$X=\begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$，β=$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$，y=$\begin{pmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{pmatrix}$; $WX = \begin{pmatrix} 2 & -4 \\ 2 & -2 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$, $Wy = \begin{pmatrix} 6 \\ 10 \\ 10 \\ 4 \\ 3 \end{pmatrix}$;

$(WX)^T WX = \begin{pmatrix} 14 & -9 \\ -9 & 25 \end{pmatrix}$, $(WX)^T Wy = \begin{pmatrix} 59 \\ -34 \end{pmatrix}$;

求解$\begin{pmatrix} 14 & -9 \\ -9 & 25 \end{pmatrix}\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 59 \\ -34 \end{pmatrix}$，得到拟合直线：

y=4.3+0.2x（未加权的拟合直线为y=4-0.1x）



内积空间应用

数据趋势分析

已知数据点

$f(t_0)$ $f(t_1)$

……

$f(t_n)$

函数趋势分析

一次型？

$y=\beta_0+\beta_1 t$

二次型？

$y=\beta_0+\beta_1 t+\beta_2 t^2$

三次型？

$y=\beta_0+\beta_1 t+\beta_2 t^2+\beta_3 t^3$

内积空间应用

数据趋势分析

趋势系数

$\hat{g}=c_0 p_0+c_1 p_1+c_2 p_2+c_3 p_3$

趋势函数

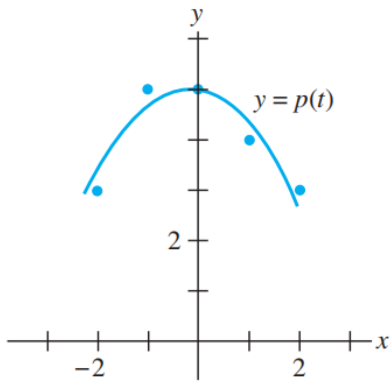
P_3空间正交基

56

内积空间应用

例题

对$(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$, $(2, 3)$进行二次型函数拟合。



内积空间应用

解析

$$
\begin{array}{lcccc}
\text{Polynomial:} & p_0 & p_1 & p_2 & \text{Data: } g \\
\text{Vector of values:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix}
\end{array}
$$

$$\Downarrow$$

$$
\begin{aligned}
\hat{p} &= \frac{\langle g, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle g, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \frac{\langle g, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 \\
&= \tfrac{20}{5} p_0 - \tfrac{1}{10} p_1 - \tfrac{7}{14} p_2
\end{aligned}
$$

$$\Downarrow$$

$$\hat{p}(t) = 4 - .1t - .5(t^2 - 2)$$

Q & A