北京大学数学科学学院
School of Mathematical Sciences
Peking University

习题课#8 (Nov. 01)

1) Schmidt 正交化 (教材 P_{152} 14题)

△ QR decomposition.

(I) Schmidt 正交化.

motivation: 设 $W=\mathrm{span}\langle\alpha_1,\cdots,\alpha_s\rangle\subseteq V=\mathbb{F}^n$. 设 $A=(\alpha_1,\cdots,\alpha_s)$.

任取向量 $\sum_i x_i\alpha_i$, $\sum_j y_j\alpha_j$ 分别记为 γ_x,γ_y.

若 $V=\mathbb{F}^n=\mathbb{R}^n$ 为 Euclidean space. 往往需研究长度、距离、夹角等信息.

长度: $\|\gamma_x\|_2^2=\|\sum_i x_i\alpha_i\|_2^2=(Ax)^T(Ax)=x^TA^TAx=x^T(A^TA)x$.

其中 $x=(x_1,\cdots,x_s)$, 对应于 γ_x 在 $\alpha_1,\cdots,\alpha_s$ 的线性组合方式.

距离: $\|\gamma_x-\gamma_y\|_2^2=\|Ax-Ay\|_2^2=(x-y)^T(A^TA)(x-y)$.

夹角: $\cos\angle(\gamma_x,\gamma_y)=\|\gamma_x\|^{-1}\|\gamma_y\|^{-1}(\gamma_x\cdot\gamma_y)=\|\gamma_x\|^{-1}\|\gamma_y\|^{-1}x^T(A^TA)y$.

自然的想法: 要使计算方便, 应当试图约定 (A^TA) 较为简单.

换句话说, 应当选取与 $\alpha_1,\cdots,\alpha_s$ 等价的另一组 $\beta_1,\cdots,\beta_t$ 使得

B^TB 尽量简单. $B=(\beta_1,\cdots,\beta_r)$. 这样一来可直接由 γ_x,γ_y

在 B 上线性组合方式 x 和 y 容易得到 γ_x,γ_y 的几何关系.

当约定 $B^TB=I_r$ 时, 有 $\beta_k^T\beta_l=\delta_{kl}$. 则 γ_x,γ_y 的几何信息

与 x,y 在 $\mathbb{R}^r$ 中的几何信息完全一致.

如何由一般的 $\alpha_1,\cdots,\alpha_s$ 出发, 得到 $\beta_1,\cdots,\beta_r$? 贪心策略. $B=\phi$.

for $i=1$ to s:

若 $\alpha_i\notin\mathrm{span}B$, 则存在某一 β_i 使得, $\|\beta_i\|=1$

$\alpha_i\in\mathrm{span}(B,\beta_i)$, 且 β_i 与 B 中所有向量均正交

将 β_i 添加至 B 中

否则 $\alpha_i\in\mathrm{span}B$. 不进行操作.

这一算法中需要考虑：① β_i是否始终存在 ② β_i如何求得.

由 $\alpha_1,\cdots,\alpha_{i-1}\in B$，$\alpha_i\notin B$ 知（联系几何直观）

$$\alpha_i=\alpha_i^{\perp}+\alpha_i''，\ \alpha_i''\in B，\ \alpha_i^{\perp}\perp \mathrm{Im}B.$$

令 $\alpha_i''=\sum_p u_p\beta_p$，则 $\alpha_i^T\beta_p=\alpha_i^{\perp}\cdot\beta_p+\alpha_i''\cdot\beta_p=u_p\beta_p^T\beta_p$.

于是 $\alpha_i''=\sum_p\frac{\alpha_i^T\beta_p}{\beta_p^T\beta_p}\beta_p$.

(II) 矩阵表示.

由Schimidt正交化过程可知，对于 $A=(\alpha_1,\cdots,\alpha_s)$ 进行正交化得到

$B=(\beta_1,\cdots,\beta_r)$，而对于每一$\beta_k$，可表示成$\alpha_k$和$\beta_1,\cdots,\beta_{k-1}$的线性组合.

利用归纳易知 β_k总能表示为$\alpha_1,\cdots,\alpha_k$的线性组合，故存在上三角阵C

满足 $B=AC$.（注意这里要求A列线性无关，否则仅能得到C为阶梯形）

由A列满秩有 $s=r$，$\mathrm{rank}A=s\leq n$. 相应地，C对角元为正，得可逆.

有 $A=BC^{-1}$. 得到关于A的QR分解.

一般地，对于 $A=(\alpha_1,\cdots,\alpha_s)$，存在上三角$\tilde{C}$使得$A=B\tilde{C}$.

对于每一α_k，由构造关系知其必能表示为$\beta_1,\cdots,\beta_k$的线性组合，

且当A列满秩时迭代过程中始终有$\alpha_i\notin \mathrm{span}B$. 于是存在上三角$C$使得

$A=BC$. 当A列满秩时可保证C的对角元全为正.

(III) 分解的唯一性.

当A列满秩时存在唯一列正交Q_1和上三角R_1，对角元为正数满足 $A=Q_1R_1$.

注意尺寸 $A\in\mathbb{R}^{n\times r}$，$Q_1\in\mathbb{R}^{n\times r}$，$R_1\in\mathbb{R}^{n\times r}$.

亦可写成正交 $Q=(Q_1,Q_2)$，$R=\begin{pmatrix}R_1\\0\end{pmatrix}$，$A=QR$，但此时

Q_2的选取将不再唯一.

Proof. 存在性已由前文给出. 下面叙述唯一性.

设 $A=Q_1R_1=\hat{Q}_1\hat{R}_1$，$\hat{Q}_1,\hat{R}_1$满足与$Q_1,R_1$相同的条件.

~~有 $\widetilde{Q}_1^T Q_1 = \widetilde{Q}_1^T Q_1 R_1 R_1^{-1} = \widetilde{R}_1 R_1^{-1}$ 为上三角阵.~~

~~而 $(\widetilde{Q}_1^T Q_1)^T(\widetilde{Q}_1^T Q_1)$~~

$$Q_1R_1 =: (\beta_1\ \hat{Q}_1)\begin{pmatrix} r_1 & * \\ & \hat{R}_1\end{pmatrix} = (\tilde{\beta}_1\ \hat{\tilde{Q}}_1)\begin{pmatrix}\tilde{r}_1 & * \\ & \hat{\tilde{R}}_1\end{pmatrix} = \widetilde{Q}_1\widetilde{R}_1$$

有 $r_1\beta_1=\tilde{r}_1\tilde{\beta}_1$. 两边同时取模: $|r_1|=|\tilde{r}_1| \Rightarrow r_1=\tilde{r}_1>0$ (均为正数)

于是 $\beta_1=\tilde{\beta}_1$. 接下来由 $\hat{Q}_1\hat{R}_1 = \hat{\tilde{Q}}_1\hat{\tilde{R}}_1$ 可依次得到 $\beta_k=\tilde{\beta}_k, \forall k$.

于是 $Q_1=\widetilde{Q}_1$, 那么 $R_1 = Q_1^TQ_1R_1 = Q_1^T\widetilde{Q}_1\widetilde{R}_1 = \widetilde{Q}_1^T\widetilde{Q}_1\widetilde{R}_1=\widetilde{R}_1$.

(IV) 最小二乘法. $A\in\mathbb{R}^{n\times m}$

在涉及求解 $Ax=b$ 的实际问题中, 往往会因为测量、计算误差使我们无法获得准确的 b, 在这种情形下可能导致方程组无解. 故往往考虑求解

$\hat{x} = \underset{x}{\arg\min}\|Ax-b\|_2 = \underset{x}{\arg\min}\|Ax-b\|_2^2$ 以下假定A列满秩.

(i) 令 $F:\mathbb{R}^m\to\mathbb{R}$, $F(x)=\|Ax-b\|_2^2$, 则 F 取到最小值的必要条件是 $\nabla F=0$.

即 $A^T(Ax-b)=0 \Rightarrow x=(A^TA)^{-1}A^Tb$. ($\operatorname{rank}A^TA=\operatorname{rank}A=m$ 满秩)

(ii) 取A的QR分解: $A=Q_1R_1$. Q_1 列正交. R_1 上三角方阵, 可逆.

$$\|Ax-b\|_2^2 = \|Q_1R_1x-b\|_2^2 = (Q_1R_1x-b)^T(Q_1R_1x-b)$$
$$=\|R_1x\|^2-2b^TQ_1R_1x+\|b\|^2$$
$$=\|R_1x-Q_1^Tb\|^2+\|b\|^2-\|Q_1^Tb\|^2$$
$$=\|R_1x-Q_1^Tb\|^2+\|b\|^2-\|Q_1^Tb\|^2$$

即最小值在 $x=R_1^{-1}Q_1^Tb$ 处取到.

注意到 $\|b\|^2-\|Q_1^Tb\|^2$ 实际为 b 到 $\operatorname{Im}A$ 的距离之平方

记 $Q_1=(\beta_1,\cdots,\beta_m)$, 则 Q_1^Tb 是 b 投影到 $\operatorname{Im}A$ 的长度平方

$(\|\sum_{k=1}^{m}\frac{b^T\beta_k}{\beta_k^T\beta_k}\beta_k\|_2^2=\sum_{k=1}^{m}(b^T\beta_k)^2$.

△ 秩不等式的总结

1) $\operatorname{rank}(A+B)\leq \operatorname{rank}A+\operatorname{rank}B$

等号成立当且仅当 A,B 同时相抵于不重叠的分块对角矩阵 $\begin{pmatrix}\hat{A} & \\ & 0\end{pmatrix},\begin{pmatrix}0 & \\ & \hat{B}\end{pmatrix}$.

2) $\operatorname{rank}(A,B)\leq \operatorname{rank}A+\operatorname{rank}B$

等号成立当且仅当 $\operatorname{Im}A\cap\operatorname{Im}B=0 \Leftarrow \operatorname{rank}(A+B)=\operatorname{rank}A+\operatorname{rank}B$

3) $\operatorname{rank}(AB)\leq \operatorname{rank}A$

等号成立当且仅当 $\ker A+\operatorname{Im}B=\mathbb{F}^n \Leftarrow B$ 行满秩

4) $\operatorname{rank}(AB)\leq \operatorname{rank}B$

等号成立当且仅当 $\ker A\cap\operatorname{Im}B=0\Leftarrow A$ 列满秩.

5) $\operatorname{rank}(AB)\geq \operatorname{rank}A+\operatorname{rank}B-n$

等号成立当且仅当 $\ker A\subseteq\operatorname{Im}B$.

6) $\operatorname{rank}(ABC)\geq \operatorname{rank}(AB)+\operatorname{rank}(BC)-\operatorname{rank}(B)$

Proof. 取 $B=P\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}Q$. $AP=(A_1\ A_2)$. $QC=\begin{pmatrix}C_1\\ C_2\end{pmatrix}$.

$\operatorname{rank}ABC=\operatorname{rank}(AP)\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}(QC)=\operatorname{rank}A_1C_1$

$\geq \operatorname{rank}A_1+\operatorname{rank}C_1-\operatorname{rank}B$

而 $\operatorname{rank}AB=\operatorname{rank}(AP)\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}Q=\operatorname{rank}(A_1,0)=\operatorname{rank}A_1$

$\operatorname{rank}BC=\operatorname{rank}P\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}(QC)=\operatorname{rank}\begin{pmatrix}C_1\\ 0\end{pmatrix}=\operatorname{rank}C_1$.

$$\begin{pmatrix}ABC & \\ & B\end{pmatrix}\longrightarrow\begin{pmatrix}ABC & BC\\ & B\end{pmatrix}\longrightarrow\begin{pmatrix}0 & BC\\ -AB & B\end{pmatrix}$$

1) $A+B=AB\Rightarrow(I-A)(I-B)=I\Rightarrow(I-B)(I-A)=I\Rightarrow A+B=BA$.

$\Downarrow$

A,B 可交换.

~~2) 令 $L(n)$, $U(n)$, $O(n)$, $SO(n)$ 分别为 n阶矩阵中~~

2) 判断下列 $\mathbb{F}^{n\times n}$ 中的子集是否为 a) $\mathbb{F}^{n\times n}$ 上的线性空间 b) 对矩阵乘积封闭，求逆封闭

① 所有上三角阵

② 所有可逆上三角阵

③ 所有对角元为正的上三角阵

④ 所有可逆矩阵

⑤ 所有正交矩阵

⑥ 所有行列式为 1 的矩阵

3) 设 $A=(A_1, A_2)\in\mathbb{R}^{n\times n}$ 则有 $(\det A)^2 \le \det(A_1^TA_1)\det(A_2^TA_2)$

进一步地，$\det\begin{pmatrix} B^TB & B^TC \\ C^TB & C^TC\end{pmatrix} \le \det(B^TB)\det(C^TC)$ $\quad B\in\mathbb{R}^{n\times s}, C\in\mathbb{R}^{n\times(n-s)}$

Proof. $\det A = \sum\limits_{u_1<\cdots<u_m}(-1)^{1+\cdots+m+u_1+\cdots+u_m} A_1\begin{pmatrix} u_1,\cdots,u_m \\ 1,\cdots,m\end{pmatrix} A\begin{pmatrix} v_{m+1},\cdots,v_n \\ m+1,\cdots,n\end{pmatrix}$

$$(\det A)^2 \le \sum_{u_1<\cdots<u_m}\left[(-1)^{u_1+\cdots+u_m}A_1\begin{pmatrix} u_1,\cdots,u_m \\ 1,\cdots,m\end{pmatrix}\right]^2 \sum_{u_1<\cdots<u_m} A_2\begin{pmatrix} v_{m+1},\cdots,v_n \\ 1,\cdots,n-m\end{pmatrix}^2$$

$$= \det(A_1^TA_1)\det(A_2^TA_2)$$

$(A=(B\ C))$

于是 $\det\begin{pmatrix} B^TB & B^TC \\ C^TB & C^TC\end{pmatrix} = \det\begin{pmatrix} B^T \\ C^T\end{pmatrix}(B\ C) = (\det A)^2 \le \det(B^TB)\det(C^TC)$

4) 一些经典平面几何的概念与结论可直接推广至 $\mathbb{R}^n$.

a) 勾股定理：$\alpha\perp\beta$ $(\alpha^T\beta=0) \Rightarrow \|\alpha+\beta\|^2 = \|\alpha\|^2+\|\beta\|^2$.

b) Cauchy 不等式：$|\alpha^T\beta| \le \|\alpha\|\cdot\|\beta\|$. 等号成立仅当 (α,β) 秩1

c) 三角不等式：$\|\alpha+\beta\| \le \|\alpha\|+\|\beta\|$.

Proof. 仅说明 b)：考虑 $F(t)=\|\alpha+t\beta\|^2=t^2\|\beta\|^2+2t\alpha^T\beta+\|\alpha\|^2$.

不妨 $\beta\neq 0$. 利用二次函数极值不难得到

$$F(t)=\left(\|\beta\|t+\frac{\alpha^T\beta}{\|\beta\|}\right)^2-\frac{(\alpha^T\beta)^2}{\|\beta\|^2}+\|\alpha\|^2.$$

而 $F(t)\geq 0,\ \forall t$. 这蕴含着 $\|\alpha\|^2\geq\|\beta\|^{-2}(\alpha^T\beta)^2$.

5) 平行四边形法则：$\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2(\|\alpha\|^2+\|\beta\|^2)$

极化等式：$\alpha^T\beta=\frac{1}{4}(\|\alpha+\beta\|^2-\|\alpha-\beta\|^2)$ in $\mathbb{R}^n$

$$\alpha^*\beta=\frac{1}{4}(\|\alpha+\beta\|^2-\|\alpha-\beta\|^2+i\|\alpha+i\beta\|^2-i\|\alpha-i\beta\|^2) \quad \text{in } \mathbb{C}^n.$$

其中 $\|\gamma\|^2:=\gamma^*\gamma=\bar{\gamma}^T\gamma$.