奇异值分解的推论

1.线性映射的范数

奇异值分解可以为$||Tv||$给出一个上界.

1.1 $||Tv||$的上界

设$T \in \mathcal{L}(V, W)$,令s_1是T的最大的奇异值,那么

$$||Tv|| \leqslant s_1||v||$$

对任意$v \in V$成立.

Proof.

令$s_1, \cdots, s_m$为T的正奇异值,令$e_1, \cdots, e_m$和$f_1, \cdots, f_m$分别为V和W中的规范正交组,且由此可以推出T的奇异值分解

$$Tv = s_1 \langle v, e_1\rangle f_1 + \cdots + s_m \langle v, e_m\rangle f_m$$

对任意的$v \in V$成立.于是

$$\begin{aligned} ||Tv||^2 &= s_1^2 \left|\langle v, e_1\rangle\right|^2 + \cdots + s_m^2 \left|\langle v, e_m\rangle\right|^2 \\ &\leqslant s_1^2 \left(\left|\langle v, e_1\rangle\right|^2 + \cdots + \left|\langle v, e_m\rangle\right|^2\right) \\ &\leqslant s_1^2||v||^2 \end{aligned}$$

不等式两侧开平方根即可证得命题.

考虑$T \in \mathcal{L}(V, W)$且s_1是T的最大奇异值.以上结果表明$||Tv|| \leqslant s_1$对所有满足$||v|| \leqslant 1$的$v \in V$成立.在上式中取$v = e_1$,可得$Te_1 = s_1 f_1$,从而$||Te_1|| = s_1$.于是我们有

$$\max\{||Tv|| : v \in V, ||v|| \leqslant 1\} = s_1$$

上述结论可以让我们直接定义T的范数,而无需借助奇异值.

1.2 定义:线性映射的范数

设$T \in \mathcal{L}(V, W)$,那么T的**范数**记作$||T||$,定义为$||T|| = \max\{||Tv|| : v \in V, ||v|| \leqslant 1\}$.

线性映射的范数和向量空间的范数有一定相似指之处,例如其具有的一些共同特征.

1.3 线性映射范数的基本性质

设$T \in \mathcal{L}(V, W)$,那么

(a) $||T|| \geqslant 0$.

(b) $||T|| = 0$当且仅当$T = \mathbf{0}$.

(c) $||\lambda T|| = |\lambda|||T||$对任意$\lambda \in \mathbb{F}$成立.

(d) $||S + T|| \leqslant ||S|| + ||T||$对所有$S \in \mathcal{L}(V, W)$成立.

这些性质是不难证明的.另外,对于任意$S, T \in \mathcal{L}(V, W)$,$||S - T||$这个量通常被称为两者的距离.不正式地说,这值越小反映了S和T越接近.另外,$||T||$还有多种等价的表达式.

1.4 $||T||$的多种表达式

设$T \in \mathcal{L}(V, W)$,那么以下式子均等于$||T||$.

(a) T的最大奇异值.

(b) $\max\{||Tv|| : v \in V, ||v|| = 1\}$.

(c) $\min\{c : \forall v \in V, ||Tv|| \leqslant c||v||\}$.

Proof.

(a) 由前面的叙述可知.

(b) 令$v \in V$且$0 < ||v|| \leqslant 1$,令$u = \dfrac{v}{||v||}$,那么有

$$||Tu|| = \left|\left|T\left(\frac{v}{||v||}\right)\right|\right| = \frac{||Tv||}{||v||} \geqslant ||Tv||$$

且$||u|| = 1$.于是关注$||Tv||$在$||v|| \leqslant 1$下的最大值只需关注所有范数为1的向量即可,于是证明了**(b)**.

(c) 设$v \in V$且$v \neq \mathbf{0}$,那么由T的定义可得

$$\left|\left|T\left(\frac{v}{||v||}\right)\right|\right| \leqslant ||T||$$

于是$||Tv|| \leqslant ||T||||v||$.

现在设$c \geqslant 0$且$||Tv|| \leqslant c||v||$对所有$v \in V$成立,那么$||Tv|| \leqslant c$对所有满足$||v|| \leqslant 1$的所有$v \in V$成立.

对于所有这样的v,上述不等式左侧取得的最大值为$||T||$,于是$||T|| \leqslant c$.于是$||T||$即是满足条件的最小数.

下面的这条结论表明线性映射和它的伴随的范数相同.

1.5 伴随的范数

设$T \in \mathcal{L}(V, W)$,那么$||T^*|| = ||T||$.

Proof.

设$w \in W$,那么

$$||T^*w||^2 = \langle T^*w, T^*w \rangle = \langle TT^*w, w \rangle \leqslant ||TT^*w||||w|| \leqslant ||T||||T^*w||||w||$$

于是

$$||T^*w|| \leqslant ||T||||w||$$

根据**1.4(c)**可知$||T^*|| \leqslant ||T||$.

将T和T^*的位置交换,可相似的证得$||T|| \leqslant ||T^*||$.于是$||T|| = ||T^*||$,命题得证.

2.用有较低维值域的线性映射进行逼近

接下来的结论是奇异值分解的重要应用.它意味着削去奇异值分解前k项之后的项,从而得到值域的维数为k的线性映射T_k是在所有值域维数至多为k的线性映射中到T的距离最小的.换句话说,在值域维数有限制的前提下,这样所得的线性映射T_k是逼近T的最佳结果.这条结果引出了压缩巨大矩阵同时尽可能保留较多信息的方法.

2.1 限制值域维数得到的线性映射的最佳逼近

设$T \in \mathcal{L}(V, W)$且$s_1, \cdots, s_m$是按序的T的正奇异值.对于任意$1 \leqslant k < m$,有

$$\min \{||T - S|| : S \in \mathcal{L}(V, W), \dim \operatorname{range} S \leqslant k\} = s_{k+1}$$

进一步地,假定T的奇异值分解为

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

令T_k由下式定义

$$T_k v = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_k \rangle f_k$$

那么T_k即是上式取等时的S.换言之,$\dim \operatorname{range} T_k = k$且$||T - T_k|| = s_{k+1}$.

Proof.

对于任意$v \in V$有

$$\begin{aligned} ||(T - T_k)v||^2 &= |(s_{k+1} \langle v, e_{k+1} \rangle f_{k+1} + \cdots + s_m \langle v, e_m \rangle f_m)| \\ &= s_{k+1}^2 |\langle v, e_{k+1} \rangle|^2 + \cdots + s_m^2 |\langle v, e_m \rangle|^2 \\ &\leqslant s_{k+1}^2 \left(|\langle v, e_{k+1} \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \right) \\ &\leqslant s_{k+1}^2 ||v||^2 \end{aligned}$$

于是$||T - T_k|| \leqslant s_{k+1}$.又因为$(T - T_k)e_{k+1} = s_{k+1} f_{k+1}$,于是$||T - T_k|| = s_{k+1}$.

设$S \in \mathcal{L}(V, W)$且$\dim \operatorname{range} S \leqslant k$,那么长度为$k+1$的组$Se_1, \cdots, Se_{k+1}$线性相关.

于是存在不全为0的$a_1,\cdots,a_{m+1}\in\mathbb{F}$使得

$$a_1Se_1+\cdots+a_{k+1}Se_{k+1}=\mathbf{0}$$

又因为$e_1,\cdots,e_{k+1}$线性无关且各a不全为0,于是$a_1e_1+\cdots+a_{m+1}e_{m+1}\neq\mathbf{0}$.我们有

$$\begin{aligned}||(T-S)\left(a_1e_1+\cdots+a_{k+1}e_{k+1}\right)||^2&=||T\left(a_1e_1+\cdots+a_{k+1}e_{k+1}\right)||^2\\&=||s_1a_1f_1+\cdots+s_{k+1}a_{k+1}f_{k+1}||^2\\&=s_1^2|a_1|^2+\cdots+s_{k+1}^2|a_{k+1}|^2\\&\geqslant s_{k+1}^2\left(|a_1|^2+\cdots+|a_{k+1}|^2\right)\\&=s_{k+1}^2||a_1e_1+\cdots+a_{k+1}e_{k+1}||^2\end{aligned}$$

于是上式意味着$||T-S||\geqslant s_{k+1}$.因此,在所有满足条件的S中,当$S=T_k$时使得$||T-S||$最小.

奇异值分解用于其它条件下求最佳拟合的线性映射的例子见习题.

3.极分解

回顾一下,我们之前讨论了复数和算子之间的类比.现在,观察下列等式

$$z=\left(\frac{z}{|z|}\right)|z|=\left(\frac{z}{|z|}\right)\sqrt{\overline{z}z}$$

根据上式,我们猜测每个$T\in\mathcal{L}(V)$都可以写成一个酉算子乘以$\sqrt{T^*T}$的结果.而上述猜测确实是正确的,它被称为极分解.

3.1 极分解

设$T\in\mathcal{L}(V)$,那么存在酉算子$S\in\mathcal{L}(V)$使得$T=S\sqrt{T^*T}$.

Proof.

考虑T的奇异值分解

$$Tv=s_1\left\langle v,e_1\right\rangle f_1+\cdots+s_m\left\langle v,e_m\right\rangle f_m$$

将$e_1,\cdots,e_m$和$f_1,\cdots,f_m$扩充为V的规范正交基$e_1,\cdots,e_n$和$f_1,\cdots,f_n$.定义$S\in\mathcal{L}(V)$为

$$Sv=\left\langle v,e_1\right\rangle f_1+\cdots+\left\langle v,e_n\right\rangle f_n$$

对任意$v\in V$成立.不难验证S是酉算子.

将T^*作用于T的奇异值分解两侧,然后考虑T^*的奇异值分解,可得

$$T^*Tv=s_1^2\left\langle v,e_1\right\rangle e_1+\cdots+s_m^2\left\langle v,e_m\right\rangle e_m$$

于是

$$\sqrt{T^*T}v=s_1\left\langle v,e_1\right\rangle e_1+\cdots+s_m\left\langle v,e_m\right\rangle e_m$$

于是

$$
\begin{aligned}
S\sqrt{T^*T}v &= S\left(s_1\left\langle v,e_1\right\rangle e_1+\cdots+s_m\left\langle v,e_m\right\rangle e_m\right)\\
&= s_1\left\langle v,e_1\right\rangle f_1+\cdots+s_m\left\langle v,e_m\right\rangle f_m\\
&= Tv
\end{aligned}
$$

于是$T=S\sqrt{T^*T}$.

我们将在习题中看到极分解中的S是最接近T的酉算子(就和$\frac{z}{|z|}$是最接近z的复数类似).

4.作用于椭球和平行体的算子

我们现在有必要研究一些向量空间中特殊的向量所构成的集合.尽管它们并非V的子空间,却有更加直观的几何形象以及研究必要.

4.1 定义:球

V中半径为1,以$\mathbf{0}$为球心的**球**,记为B,定义为$B=\{v\in V:||v||\leqslant 1\}$.

特别地,在维度为2的向量空间中,我们有时候会用圆盘代替球这一名称,以免引起混乱.下面的一些定义,在二维空间中亦有相似的情形.

4.2 定义:椭球

设$f_1,\cdots,f_n$为V的规范正交基,$s_1,\cdots,s_n\in\mathbb{R}_+$.**主轴**为$s_1f_1,\cdots,s_nf_n$的**椭球**$E(s_1f_1,\cdots,s_nf_n)$定义为

$$
E(s_1f_1,\cdots,s_nf_n)=\left\{v\in V:\frac{\left|\left\langle v,f_1\right\rangle\right|^2}{s_1^2}+\cdots+\frac{\left|\left\langle v,f_n\right\rangle\right|^2}{s_n^2}\leqslant 1\right\}
$$

根据帕塞瓦尔恒等式,对于任意V的规范正交基$f_1,\cdots,f_n$,$E(f_1,\cdots,f_n)$都是V中的球.

我们现在引入一个新的记号,以体现这些具有几何直观的集合在映射下的变化.

4.3 定义:$T(\Omega)$

对于定义在V上的函数T以及$\Omega\subseteq V$,定义$T(\Omega)$为$T(\Omega)=\{Tv:v\in\Omega\}$.

根据上述定义,不难看出$T(V)=\text{range }T$.从前,我们只能大致从基向量的变化理解线性映射的几何直观.下面的定理给出了另一种更为符合直觉的方式.

4.4 可逆算子化球为椭球

设$T\in\mathcal{L}(V)$是可逆的,那么T将V中的球B映成V中的椭球.

Proof.

考虑T的奇异值分解

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

由于T是可逆的,因而$n = \dim V$,于是$e_1, \cdots, e_n$和$f_1, \cdots, f_n$是V的规范正交基.于是我们有

$$\begin{aligned} v \in B \Leftrightarrow ||v|| \leqslant 1 \Leftrightarrow ||v||^2 \leqslant 1 \\ \Leftrightarrow |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 \leqslant 1 \\ \Leftrightarrow \frac{|\langle Tv, f_1 \rangle|^2}{s_1^2} + \cdots + \frac{|\langle Tv, f_n \rangle|^2}{s_n^2} \leqslant 1 \\ \Leftrightarrow Tv \in E(s_1 f_1, \cdots, s_n f_n) \end{aligned}$$

从而$T(B) = E(s_1 f_1, \cdots, s_n f_n)$.

回顾一下,对于$u \in V, \Omega \subseteq V$,我们用$u + \Omega$表示$\Omega$的平移.

4.5 定义:$P(v_1, \cdots, v_n)$,平行体

设$v_1, \cdots, v_n$是V的一组基,令

$$P(v_1, \cdots, v_n) = \{a_1 v_1 + \cdots + a_n v_n : a_1, \cdots, a_n \in (0, 1)\}$$

平行体是形如$u + P(v_1, \cdots, v_n)$的集合,其中$u \in V$.$v_1, \cdots, v_n$被称为此平行体的**边**.

同样的,在二维的情形下我们更倾向于用平行四边形形容这样的集合.我们有如下定理.

4.6 可逆算子化平行体为平行体

设$u \in V$且$v_1, \cdots, v_n$为V的基,那么对于任意$T \in \mathcal{L}(V)$,如果T可逆,那么

$$T\left(u + P(v_1, \cdots, v_n)\right) = Tu + P(Tv_1, \cdots, Tv_n)$$

上述命题的证明是十分简单的,在此就略去.我们还可以考虑V中一类特殊的平行体:长方体(在二维空间中即矩形).

4.7 定义:长方体

V中的**长方体**是形如

$$u + P(r_1 e_1 + \cdots + r_n e_n)$$

的集合,其中$u \in V$,$r_1, \cdots, r_n \in \mathbb{R}_+$,$e_1, \cdots, e_n$是$V$的规范正交基.

考虑**4.6**的更特殊的形式,即可逆算子T是否能将长方体映至长方体?这在一般的情形下是显然不成立的,但我们能找到一些特殊的长方体使其符合条件.

4.8 每个可逆算子都将某些长方体化成长方体

设可逆算子$T \in \mathcal{L}(V)$的奇异值分解

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

其中$e_1, \cdots, e_n$和$f_1, \cdots, f_n$均为V的规范正交基.那么,对于任意$u \in V$和$r_1, \cdots, r_n \in \mathbb{R}_+$,$T$将长方体$u + P(r_1 e_1, \cdots, r_n e_n)$映成长方体$Tu + P(r_1 s_1 f_1, \cdots, r_n s_n f_n)$.

这命题的证明也是简单的,一样略去.

我们接下来讨论如何通过奇异值计算体积.由于体积是分析学的内容,于是我们在线性代数中只需直观地理解即可.具体来说,令长方体$u + P(r_1 e_1 + \cdots + r_n e_n)$的体积为$r_1 \cdots r_n$(这符合我们在二维和三位空间中对体积的定义).考虑**4.8**中的长方体,可知这样的T将体积为$r_1 \cdots r_n$的长方体映成体积为$r_1 \cdots r_n \cdot s_1 \cdots s_n$的长方体.

我们用分析学中的逼近的思想考虑V中的椭球,将其用这样的长方体近似代替后取极限可知T也将椭球的体积变化为原来的$s_1 \cdots s_n$倍.于是我们不加严格证明地给出如下结论.

4.9 体积变化倍数为奇异值的乘积

设$\mathbb{F} = \mathbb{R}$,$T \in \mathcal{L}(V)$可逆,且(可测的)$\Omega \subseteq V$.记volumn Ω表示Ω的体积,$s_1, \cdots, s_n$为T的奇异值,那么

$$\text{volumn } T(\Omega) = \prod_{k=1}^{n} s_k \cdot \text{volumn } \Omega$$

当我们学习了行列式时,就会发现$|\det T|$恰好等于T的奇异值的乘积.