6.4 正交投影和极小化问题

Tiao Lu

Peking University

2019.11.22

Gram-Schmidt 正交化

补讲 Gram-Schmidt 正交化, 不用下载 letureLA13.pdf, 相关部分已经放在 lectureLA13b.pdf 中

Gram-Schmidt 正交化

算子的规范正交上三角化

坐标变换

正交投影和极小化问题

作业

Gram-Schmidt 正交化

下面讲如何从一个线性无关的向量组得到一个规范正交的向量组。

Gram-Schmidt procedure (格拉姆-施密特过程)

Orthonormal basis 很有用? 那么是不是每个 finite dimensional 内积 space 都有一个 orthonormal basis 呢?

Gram-Schmidt procedure (格拉姆-施密特过程)

Orthonormal basis 很有用？那么是不是每个 finite dimensional 内积 space 都有一个 orthonormal basis 呢？ YES. You will see why and how.

我们所用的方法，是把一个线性无关的向量组通过一个算法 (Gram-Schmidt procedure) 来转化成一个具有相同张成的 orthonormal vector group
规范正交向量组

Gram-Schmidt procedure (格拉姆-施密特过程)

Orthonormal basis 很有用? 那么是不是每个 finite dimensional 内积 space 都有一个 orthonormal basis 呢? YES. You will see why and how.

我们所用的方法, 是把一个线性无关的向量组通过一个算法 (Gram-Schmidt procedure) 转化成一个具有相同张成的 orthonormal vector group.

规范正交向量组

Remark: 两个向量组具有同样的张成 (span)

中西结合表达更清楚 | Transform a 线性无关 vector 组 into a orthonormal 向量组 with the same span

Gram-Schmidt procedure

6.20 Gram-Schmidt procedure：如果 $(v_1, \ldots, v_m)$ 是 V 中的一个线性无关组，则 V 中有一个规范正交向量组 $(e_1, e_2, \ldots, e_m)$ 使得

6.21 $$\operatorname{span}(v_1, \ldots, v_j) = \operatorname{span}(e_1, \ldots, e_j), \quad j = 1, \ldots, m.$$

Gram-Schmidt procedure

6.20 Gram-Schmidt procedure：如果 $(v_1, \dots, v_m)$ 是 V 中的一个线性无关组，则 V 中有一个规范正交向量组 $(e_1, e_2, \dots, e_m)$ 使得

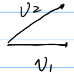
6.21 $$\operatorname{span}(v_1, \dots, v_j) = \operatorname{span}(e_1, \dots, e_j), \quad j = 1, \dots, m.$$

Remark: $\operatorname{span}(v_1) = \operatorname{span}(e_1)$ 可知 v_1 和 e_1 是线性相关的

$$\operatorname{span}(v_1, v_2) = \operatorname{span}(e_1, e_2)$$
$$\vdots$$
$$\operatorname{span}(v_1, \dots, v_m) = \operatorname{span}(e_1, e_2, \dots, e_m).$$

例: $\mathbb{R}^2$ 中的 $(\overset{v_1}{(1,0)}, \overset{v_2}{(1,1)})$. 可取 $e_1 = v_1$.

$\operatorname{span}(e_1, v_2) = \operatorname{span}(v_1, v_2)$



Gram-Schmidt procedure

6.20 Gram-Schmidt procedure：如果 $(v_1,\cdots,v_m)$ 是 V 中的一个线性无关组，则 V 中有一个规范正交向量组 $(e_1,e_2,\cdots,e_m)$ 使得

6.21 $$\operatorname{span}(v_1,\cdots,v_j)=\operatorname{span}(e_1,\cdots,e_j),\ j=1,\cdots,m.$$

Remark: $\operatorname{span}(v_1)=\operatorname{span}(e_1)$ 可知 v_1 和 e_1 是线性相关的

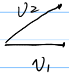
$$\operatorname{span}(v_1,v_2)=\operatorname{span}(e_1,e_2)$$

$$\vdots$$

$$\operatorname{span}(v_1,\cdots,v_m)=\operatorname{span}(e_1,e_2,\cdots,e_m)$$

例: $\mathbb{R}^2$ 中的 $(\overset{v_1}{(1,0)},\ \overset{v_2}{(1,1)})$，可取 $e_1=v_1$.

$\operatorname{span}(e_1,v_2)=\operatorname{span}(v_1,v_2)$



$\operatorname{span}(e_1,e_2)=\operatorname{span}(e_1,v_2)$?

如何 得到 e_2，使得 $e_2\perp e_1$ 且

$v_2 \in \mathrm{span}(e_1, e_2)$

那么 $v_2 = c_1 e_1 + c_2 e_2$ ①

①式两边同时和 e_1 作内积，可得

$\langle v_2, e_1 \rangle = \langle c_1 e_1 + c_2 e_2, e_1 \rangle$ ② 代入

因为 (e_1, e_2) 是 orthonormal group 所以 $\boxed{\begin{aligned}\langle e_1, e_2 \rangle &= 0 \\ \langle e_1, e_1 \rangle &= 1\end{aligned}}$

可得 $c_1 = \langle v_2, e_1 \rangle$

$v_2 = \langle v_2, e_1 \rangle e_1 + c_2 e_2$

这称作 v_2 在 e_1 上的投影，记作 $\mathrm{proj}_{e_1}(v_2)$

$$v_2 \in \mathrm{span}(e_1, e_2)$$

那么 $\quad v_2 = c_1 e_1 + c_2 e_2 \quad ①$

①式两边同时和 e_1 作内积，得

$$\langle v_2, e_1 \rangle = \langle c_1 e_1 + c_2 e_2, e_1 \rangle \quad ②$$

代入

因为 (e_1, e_2) 是 orthonormal group 所以 $\langle e_1, e_2 \rangle = 0$，$\langle e_1, e_1 \rangle = 1$

得 $\quad c_1 = \langle v_2, e_1 \rangle$

$$v_2 = \langle v_2, e_1 \rangle e_1 + c_2 e_2$$

这称作 v_2 在 e_1 上的投影，记作 $\mathrm{proj}_{e_1}(v_2)$

其中的 $c_2 e_2$ 如何确定呢？

$$v_2 \in \mathrm{span}(e_1, e_2)$$

那么 $\quad v_2 = c_1 e_1 + c_2 e_2 \quad$ ①

①式两边同时和 e_1 作内积，可得

$$\langle v_2, e_1 \rangle = \langle c_1 e_1 + c_2 e_2, e_1 \rangle \quad ②$$

代入

因为 (e_1, e_2) 是 orthonormal group 所以 $\boxed{\begin{matrix}\langle e_1, e_2 \rangle = 0 \\ \langle e_1, e_1 \rangle = 1\end{matrix}}$

可得 $\quad c_1 = \langle v_2, e_1 \rangle$

$$v_2 = \underbrace{\langle v_2, e_1 \rangle e_1}_{} + c_2 e_2$$

这称作 v_2 在 e_1 上的投影，记作 $\mathrm{proj}_{e_1}(v_2)$

其中的 $c_2 e_2$ 如何确定呢？

首先，$c_2 \neq 0$ 不然 v_2 和 e_1 是线性相关的，也即 v_2 和 v_1 是线相关的，矛盾。

$v_2 \in \mathrm{span}(e_1, e_2)$

那么 $\quad v_2 = c_1 e_1 + c_2 e_2 \quad$ ①

①式两边同时和 e_1 作内积，得

$$\langle v_2, e_1 \rangle = \langle c_1 e_1 + c_2 e_2, e_1 \rangle \quad ②$$

代入

因为 (e_1, e_2) 是 orthonormal group 所以 $\boxed{\begin{array}{l}\langle e_1, e_2 \rangle = 0 \\ \langle e_1, e_1 \rangle = 1\end{array}}$

得 $\quad c_1 = \langle v_2, e_1 \rangle$

$$v_2 = \langle v_2, e_1 \rangle e_1 + c_2 e_2$$

这称作 v_2 在 e_1 上的投影，记作 $\mathrm{proj}_{e_1}(v_2)$

其中的 $c_2 e_2$ 如何确定呢？

首先，$c_2 \neq 0$ 不然 v_2 和 e_1 是线性相关的，但前面 v_2 和 v_1 是线性无关的，矛盾。这样，$e_2 = \frac{1}{c_2}(v_2 - \mathrm{proj}_{e_1}(v_2))$ 下页继续

接上页.

e_2 和 $\underbrace{v_2 - proj_{e_1}(v_2)}$

v₂在e₁上的投影

仅仅相差一个标量倍，而且 $\|e_2\| = 1$.

这样，我们只要把 $v_2 - proj_{e_1}(v_2)$ 单位化就得到 e_2

所以

$$e_2 = \pm \frac{v_2 - proj_{e_1}(v_2)}{\|v_2 - proj_{e_1}(v_2)\|}$$

接上页.

e_2 和 $v_2 - \underbrace{proj_{e_1}(v_2)}_{v_2\text{在}e_1\text{上的投影}}$

仅仅相差一个标量倍，而且 $\|e_2\| = 1$.

这样，我们只要把 $v_2 - proj_{e_1}(v_2)$ 单位化就得到 e_2

所以

$$e_2 = \underbrace{\pm}_{\text{一般取正号}} \frac{v_2 - proj_{e_1}(v_2)}{\| v_2 - proj_{e_1}(v_2) \|}$$

再验证 $span(e_1, e_2) = span(v_1, v_2)$.

接上页.

$$e_2 \text{ 和 } v_2 - \underbrace{\text{proj}_{e_1}(v_2)}_{v_2\text{在}e_1\text{上的投影}}$$

仅仅相差一个标量倍，而且 $\|e_2\| = 1$.

这样，我们只要把 $v_2 - \text{proj}_{e_1}(v_2)$ 单位化就得到 e_2

所以
$$e_2 = \frac{v_2 - \text{proj}_{e_1}(v_2)}{\underbrace{\|v_2 - \text{proj}_{e_1}(v_2)\|}_{c_2}} \quad (1)$$

再验证 $\text{span}(e_1, e_2) = \text{span}(v_1, v_2)$.

显然 $\text{span}(e_1, e_2) \subset \text{span}(e_1, v_2) = \text{span}(v_1, v_2)$

再利用 (1) 可得 $v_2 = c_2 e_2 + \underbrace{\text{proj}_{e_1}(v_2)}_{\text{这是}v_2\text{在}e_1\text{上的投影}}$

接上页.

e_2 和 $v_2 - \underbrace{proj_{e_1}(v_2)}_{v_2\text{在}e_1\text{上的投影}}$

仅仅相差一个标量倍，而且 $\|e_2\|=1$.

这样.我们只要把 $v_2 - proj_{e_1}(v_2)$ 单位化就得到 e_2

所以

$$e_2 = \frac{v_2 - proj_{e_1}(v_2)}{\underbrace{\|v_2 - proj_{e_1}(v_2)\|}_{c_2}} \quad ①$$

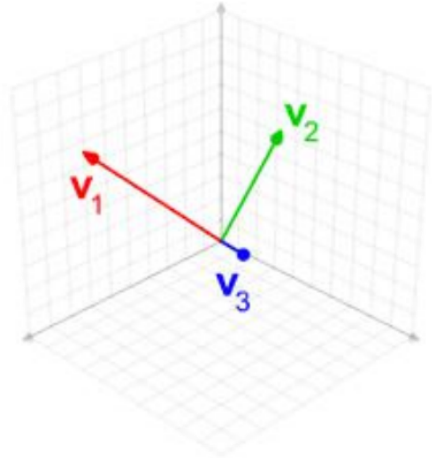
再验证 $span(e_1, e_2) = span(v_1, v_2)$ ②

显然 $span(e_1, e_2) \subset span(e_1, v_2) = span(v_1, v_2)$

再利用 ① 可得 $v_2 = c_2 e_2 + proj_{e_1}(v_2) \in span(e_1, e_2)$

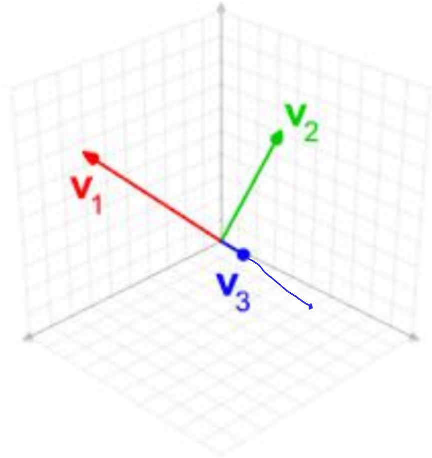
故 $span(v_1, v_2) \subset span(e_1, e_2)$ ③ . 由 ② ③ 就得证 $span(e_1, e_2) = span(v_1, v_2)$.

R^3 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程



v_3 画得比较短

R^3 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程

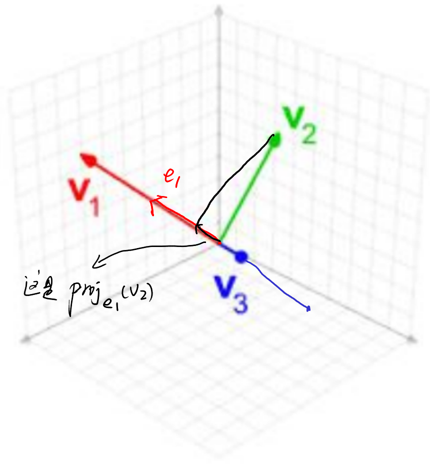


v_3 画得比较短（划长点）

$u_1 = u_1$

$e_1 = \dfrac{u_1}{\|u_1\|}$ 单位化

R^3 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程



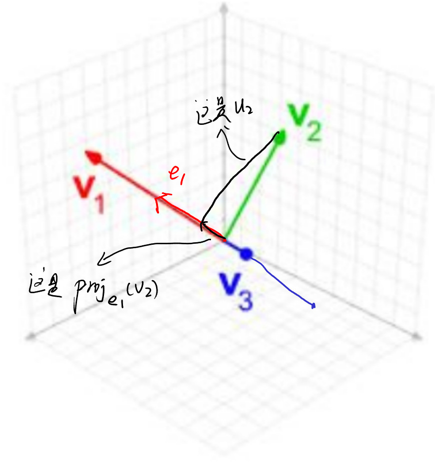
v_3 画得比较短 (训练点)

$u_1 = u_1$

$e_1 = \frac{u_1}{\|u_1\|}$ 单位化

$u_2 = v_2 - proj_{e_1}(v_2)$

$\mathbb{R}^3$ 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程



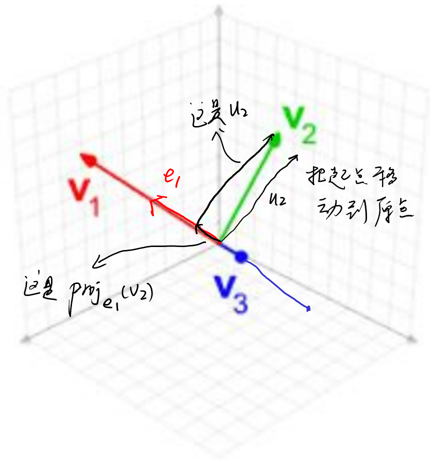
v_3 画得比较短（别惦点）

$u_1 = v_1$

$e_1 = \dfrac{u_1}{\|u_1\|}$ 单位化

$u_2 = v_2 - proj_{e_1}(v_2)$

R^3 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程



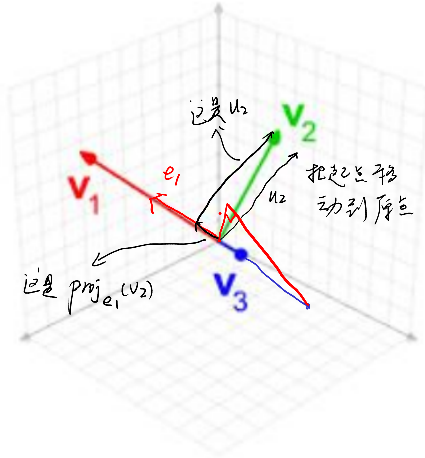
v_3 画得比较短（训快点）

$$u_1 = u_1$$

$$e_1 = \frac{u_1}{\|u_1\|} \quad \text{单位化}$$

$$u_2 = v_2 - \mathrm{proj}_{e_1}(v_2)$$

R^3 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程



v_3 画得比较短（别忙点）

$u_1 = u_1$

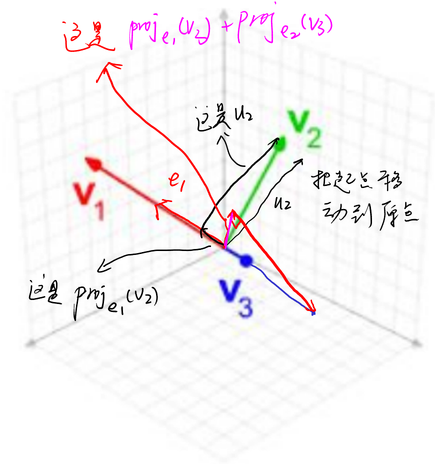
$e_1 = \frac{u_1}{\|u_1\|}$ 单位化

$u_2 = v_2 - proj_{e_1}(v_2)$ 正交化

$e_2 = \frac{u_2}{\|u_2\|}$ 单位化

$u_3 = v_3 - proj_{e_1}(v_3) - proj_{e_2}(v_3)$

$\mathbb{R}^3$ 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程



v_3 画得比较短（划快点）

$u_1 = u_1$

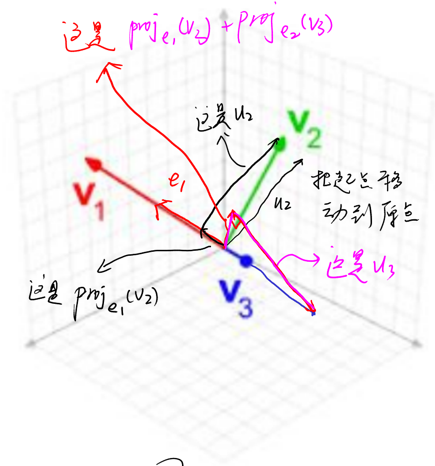
$e_1 = \frac{u_1}{\|u_1\|}$ 单位化

$u_2 = v_2 - proj_{e_1}(v_2)$ 正交化

$e_2 = \frac{u_2}{\|u_2\|}$ 单位化

$u_3 = v_3 - proj_{e_1}(v_3) - proj_{e_2}(v_3)$

R^3 中的 (v_1, v_2, v_3) 的 Gram-Schmidt 正交化过程



v_3 画得的较短（到此点）

$$u_1 = v_1$$

$$e_1 = \frac{u_1}{\|u_1\|} \quad \text{单位化}$$

$$u_2 = v_2 - proj_{e_1}(v_2) \quad \text{正交化}$$

$$e_2 = \frac{u_2}{\|u_2\|} \quad \text{单位化}$$

$$u_3 = v_3 - proj_{e_1}(v_3) - proj_{e_2}(v_3)$$

$$e_3 = \frac{u_3}{\|u_3\|} \quad \text{单位化}$$

图示 Gram-Schmidt 正交化过程

Gram-Schmidt procedure

6.20 格拉姆–施密特过程 (Gram-Schmidt procedure)： 如果 $(v_1,\cdots v_m)$ 是 V 中的线性无关向量组，则 V 有规范正交向量组 $(e_1,\cdots,e_m)$ 使得

6.21 $\operatorname{span}(v_1,\cdots,v_j)=\operatorname{span}(e_1,\cdots,e_j),\quad j=1,\cdots,m.$

证明：

$$u_1 = v_1 \qquad e_1 = \frac{u_1}{\|u_1\|}$$

$$u_2 = v_2 - \operatorname{proj}_{e_1}(v_2) \qquad e_2 = \frac{u_2}{\|u_2\|}$$

$$u_3 = v_3 - \operatorname{proj}_{e_1}(v_3) - \operatorname{proj}_{e_2}(v_3) \qquad e_3 = \frac{u_3}{\|u_3\|}$$

$$\vdots$$

$$u_m = v_m - \sum_{j=1}^{m-1} \operatorname{proj}_{e_j}(u_m) \qquad e_m = \frac{u_m}{\|u_m\|}.$$

you do it.

验证 $(e_1,\cdots,e_m)$ 是规范正交向量组 (you can do it!)

再验证 $\operatorname{span}(e_1,\cdots,e_j) = \operatorname{span}(v_1,\cdots,v_j)$, $j=1,2,\cdots,m$. □

每个 finite dimensional inner produce space 都有 orthonomal basis

6.24 推论：每个有限维内积空间都有规范正交基.

证明：V is a finite dimensional space. so there is a basis $(v_1, \cdots, v_n)$，应用上面的 Gram-schmidt 正交化过程. 我们得到一个规范正交的向量组 $(e_1, \cdots, e_n)$

而且 $\text{span}(e_1, \cdots, e_n) = \text{span}(v_1, \cdots, v_n) = V$

且 $(e_1, \cdots, e_n)$的长度是 n.

所以 $(e_1, \cdots, e_n)$ 是 V 的一个 orthonormal basis.

每个 finite dimensional inner produce space 都有 orthonomal basis

6.24 推论：每个有限维内积空间都有规范正交基.

证明：V is a finite dimensional space. so there is a basis $(v_1, \cdots, v_n)$，应用上面的 Gram-schmidt 正交化过程. 我们得到一个规范正交的向量组 $(e_1, \cdots, e_n)$

而且 $\operatorname{span}(e_1, \cdots, e_n) = \operatorname{span}(v_1, \cdots, v_n) = V$

且 $(e_1, \cdots, e_n)$ 的长度是 n.

所以 $(e_1, \cdots, e_n)$ 是 V 的一个 orthonormal basis.

6.25 推论：V 中的每个规范正交的向量组都可以扩充成 V 的一个规范正交基。

证明：$(e_1, \cdots, e_m)$ 是 V 的一个正交组. 必然是线性无关的. 下页继续

从命 on 扩充成 V 的一个基 $(e_1, \dots, e_m, f_1, \dots, f_n)$

然后对这个基应用 Gram-Schmidt 正交化过程.

就得到 V 的一个规范正交基

$$(\underbrace{e_1, \dots, e_m}, e_{m+1}, \dots, e_{m+n})$$

为啥 Gram-Schmidt procedure 没有改变 $(e_1, \dots, e_m)$?

原因是 …… 你自己做一遍看看就知道了.

从命ann扩充成 V 的一个基 $(e_1, \dots, e_m, f_1, \dots, f_n)$

然后对这个基应用 Gram-Schmidt 正交化过程.

就得到 V 的一个规范正交基

$$(\underbrace{e_1, \dots, e_m}, e_{m+1}, \dots, e_{m+n})$$

为啥 Gram-Schmidt procedure 没有改变 $(e_1, \dots, e_m)$?

Gram-Schmidt 正交化

算子的规范正交上三角化

坐标变换

正交投影和极小化问题

作业

算子的规范正交上三角化

一个可上三角化的算子一定可以规范正交上三角化。

回想一下, 如果一个矩阵对角线下方的所有元素都等于 0, 则称这个矩阵是上三角的. 也就是说, 一个上三角矩阵形如:

$$\begin{bmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{bmatrix}.$$

在上一章我们证明了, 如果 V 是复向量空间, 那么对 V 上的每个算子都存在一个基使得该算子关于此基具有上三角矩阵 (参见 5.13). 既然现在讨论的是内积空间, 我们想知道何时存在规范正交基使得算子关于此基有上三角矩阵. 下一个推论表明, 能使得 T 具有上三角矩阵的基的存在性蕴含着具有同样性质的规范正交基的存在性. 这一结果在实向量空间和复向量空间上都成立 (但在实向量空间上, 这个假设只对某些算子成立).

6.27 推论：假设 $T \in \mathcal{L}(V)$. 如果 T 关于 V 的某个基具有上三角矩阵，那么 T 关于 V 的某个规范正交基也具有上三角矩阵.

证明：设 T 关于 V 的基 $(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n)$ 具有上三角矩阵，那么对每个 $j = 1, \cdots, n$，都有 $\mathrm{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_j)$ 在 T 下是不变的（参见 5.12）.

对 $(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n)$ 应用格拉姆-施密特过程，得到了 V 的规范正交基 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)$. 因为对每个 j 都有

$$\mathrm{span}(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_j) = \mathrm{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_j)$$

（参见 6.21），所以，对每个 $j = 1, \cdots, n$，$\mathrm{span}(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_j)$ 在 T 下都是不变的. 因此，由 5.12，T 关于规范正交基 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)$ 具有上三角矩阵. ■

下一个结果是上一个推论的重要应用.

6.28 推论： 设 V 是复向量空间，并且 $T \in \mathcal{L}(V)$，则 T 关于 V 的某个规范正交基具有上三角矩阵.

证明： 由 5.13 和 6.27 立即可得. ■

这个结果有时称为舒尔定理. 德国数学家舒尔 (Issai Schur) 在 1909 年发表了这个结果的第一个证明.

Gram-Schmidt 正交化

算子的规范正交上三角化

坐标变换

正交投影和极小化问题

作业

坐标变换

设 $(e_1, \cdots, e_n)$ 和 $(f_1, \cdots, f_n)$ 是线性空间 V 的两个基, 如果

$$\begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix} = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} \begin{bmatrix} s_{1,1} & \cdots & s_{1,n} \\ \vdots & \ddots & \vdots \\ s_{n,1} & \cdots & s_{n,n} \end{bmatrix}$$

记上面的 $n \times n$ 为矩阵 $S = [s_{i,j}]$. 那么 (证明我放在下一页)

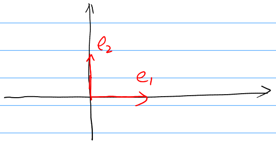
$$\mathcal{M}(T, (f_1, \cdots, f_n)) = S^{-1} \mathcal{M}(T, (e_1, \cdots, e_n)) S.$$

其中 S^{-1} 是 S 的逆矩阵, 并且显然有

$$\begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} = \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix} S^{-1}$$

可逆的线性把一个基映成一个基

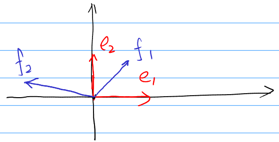
(e_1, e_2) 是 $\mathbb{R}^2$ 的标准基，



线性映射 $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$Te_1 = f_1$$
$$Te_2 = f_2$$

(e_1, e_2) 是 $\mathbb{R}^2$ 的标准基,



线性映射 $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by

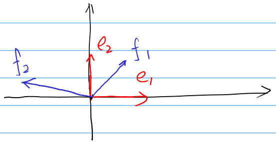
$$Te_1 = f_1$$
$$Te_2 = f_2$$

You can prove that T is invertible $\Leftrightarrow$ (f_1, f_2) is a basis

T 的矩阵是什么呢? $M(T, (e_1, e_2), (e_1, e_2)) = \begin{array}{c|cc} & Te_1 & Te_2 \\ \hline e_1 & a_{11} & a_{12} \\ e_2 & a_{21} & a_{22} \end{array}$

由映射矩阵的定义知 $\rightarrow \begin{cases} Te_1 = a_{11} e_1 + a_{21} e_2 \\ Te_2 = a_{12} e_1 + a_{22} e_2 \end{cases}$

(e_1, e_2) 是 $\mathbb{R}^2$ 的标准基，



线性映射 $T: \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$Te_1 = f_1$$
$$Te_2 = f_2$$

You can prove that T is invertible $\Leftrightarrow$ (f_1, f_2) is a basis

T 的矩阵是什么呢？ $M(T, (e_1, e_2), (e_1, e_2)) = \begin{array}{c} \\ e_1 \\ e_2 \end{array} \begin{array}{c} \begin{array}{cc} Te_1 & Te_2 \end{array} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{array}$

由映射矩阵的定义知

$$\begin{cases} Te_1 = a_{11} e_1 + a_{21} e_2 = f_1 \\ Te_2 = a_{12} e_1 + a_{22} e_2 = f_2 \end{cases}$$

$$\begin{cases} f_1 = a_{11} e_1 + a_{21} e_2 \\ f_2 = a_{12} e_1 + a_{22} e_2 \end{cases}$$

写成

$$[f_1, f_2] = [e_1, e_2] \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

$$\begin{cases} f_1 = a_{11}e_1 + a_{21}e_2 \\ f_2 = a_{12}e_1 + a_{22}e_2 \end{cases}$$

写成

$$[f_1, f_2] = [e_1, e_2] \underbrace{\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}}_{\text{是 } \mathbb{R}^{2\times 2} \text{ 的矩阵}}$$

这个写法是和我们以前学习的知识不相容的，

我们仅仅学习两个矩阵的乘法. $A \in F^{s\times n}$, $B \in F^{s\times m}$

$$\underbrace{(AB)_{i,j}}_{(AB)_{i,j}\text{ 表示 } AB \text{ 的 } (i,j) \text{ 元}} = \sum_{k=1}^{n} \underbrace{A_{i,k}}_{A_{i,k}\text{ 表示 } A \text{ 的 } (i,k) \text{ 元}} B_{k,j}$$

$$\begin{cases} f_1 = a_{11} e_1 + a_{21} e_2 \\ f_2 = a_{12} e_1 + a_{22} e_2 \end{cases}$$

写成

$$[f_1, f_2] = \boxed{[e_1, e_2] \underbrace{\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}}_{\text{是 } \mathbb{R}^{2\times 2} \text{ 的矩阵}}}$$

这个写法是和我们以前学习的知识不相容的，

我们仅仅学习两个矩阵的乘法. $A \in F^{s\times n}$, $B \in F^{s\times m}$

$$\underbrace{(AB)_{i,j}} = \sum_{k=1}^{n} \underbrace{A_{i,k}} B_{k,j}$$

$(AB)_{i,j}$ 表示 AB 的 (i,j) 元　　$A_{i,k}$ 表示 A 的 (i,k) 元

这不是两个矩阵的相乘

$$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

右端仅仅是形式上看成了两矩阵相乘，得

$$[f_1, f_2] = \boxed{[e_1, e_2]} \begin{bmatrix} \boxed{\begin{matrix} a_{11} \\ a_{21} \end{matrix}} & \begin{matrix} a_{12} \\ a_{22} \end{matrix} \end{bmatrix}$$

右端仅仅是形式上看成了两矩阵相乘，该

$$e_1 a_{11} + e_2 a_{21}$$

$$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{bmatrix}$$

右端仅仅是形式上看成了两个矩阵相乘，得

$$\underbrace{e_1 a_{11} + e_2 a_{21}}$$

这实际上又带了一个新问题

$e_1 a_{11}$ 是没有定义过的运算，因为我们仅仅定义过

向量乘数

数乘向量

$$[f_1, f_2] = [\boxed{e_1, e_2}] \begin{bmatrix} \boxed{\begin{matrix} a_{11} \\ a_{21} \end{matrix}} & \begin{matrix} a_{12} \\ a_{22} \end{matrix} \end{bmatrix}$$

右端仅仅是形式上看成了两个矩阵相乘，

且把向量乘数

变换为数乘向量。

$e_1 a_{11} + e_2 a_{21}$

这实际上又带了一个新问题

可得

$[f_1, f_2] = [a_{11}e_1 + a_{21}e_2, a_{12}e_1 + a_{22}e_2]$

$e_1 a_{11}$ 是没有定义过的运算，因为我们

向量乘 数

仅仅定义过 数乘向量

解决方法是把 $e_1 a_{11}$ 理解成 $a_{11} e_1$

$a_{11} e_1 + a_{21} e_2$

新基　　　老基　　　过渡矩阵 (transition matrix)

$$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

右端仅仅是形式上看成了两矩阵相乘，

且把向量乘数

变换为数乘向量。

$e_1 a_{11} + e_2 a_{21}$

这实际上又带了一个新问题

得

$[f_1, f_2] = [a_{11}e_1 + a_{21}e_2,\ a_{12}e_1 + a_{22}e_2]$

$e_1 a_{11}$ 是没有定义过的运算，因为我

向量乘数

们仅仅定义过数乘向量

解决方法是把 $e_1 a_{11}$ 理解成 $a_{11}e_1$

$a_{11}e_1 + a_{21}e_2$

新基　　　　老基　　　　过渡矩阵 (transition matrix)

$$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

右端仅仅是形式上看成了两矩阵相乘，

且把向量乘数

变换为数乘向量。

$e_1 a_{11} + e_2 a_{21}$

cons:
① 向量乘数
不自然

这实际上又带了一个新问题

得

$$[f_1, f_2] = [a_{11}e_1 + a_{21}e_2,\ a_{12}e_1 + a_{22}e_2]$$

$e_1 a_{11}$ 是没有定义过的运算，因为我们

向量乘数

仅仅定义过
数乘向量

解决方法是把 $e_1 a_{11}$ 理解成 $a_{11} e_1$

$a_{11} e_1 + a_{21} e_2$

新基 老基 过渡矩阵 (transition matrix)

$$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

cons:
① 向量乘数
不自然
pros:
① Transition matrix $= M(T; (e_1, e_2))$

右端仅仅是形式上看成了"矩阵相乘",

且把向量乘数

变换为数乘向量.

$e_1 a_{11} + e_2 a_{21}$

这实际上又带了一个新问题

得

$[f_1, f_2] = [a_{11}e_1 + a_{21}e_2, a_{12}e_1 + a_{22}e_2]$

$e_1 a_{11}$ 是没有定义过的运算, 因为我们

向量乘数

仅仅定义过 数乘向量

↓解决方法是把 $e_1 a_{11}$ 理解成 $a_{11}e_1$

$a_{11}e_1 + a_{21}e_2$

另一种方法来写坐标变换

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$[f_1, f_2] = [e_1, e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

共同点

都不是矩阵相乘，却利用矩阵相乘的法则来算

另一种方法来写坐标变换

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

	$[f_1, f_2] = [e_1, e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$	$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$
共同点	都不是矩阵相乘，却利用矩阵相乘的法则来算	
数乘向量 向量乘数	向量乘数 不自然	数乘向量 自然

另一种方法来写坐标变换

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

	$[f_1, f_2] = [e_1, e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$	$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$
共同点	都不是矩阵相乘，却利用矩阵相乘的法则来算	
数乘向量 向量乘数	向量乘数 不自然	数乘向量 自然
过渡矩阵 和 $f_i = Te_i$ 的 $M(T)$	过渡矩阵 $= M(T;(e_1,e_2))$	

另一种方法来写坐标变换

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

	$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$	$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$
共同点	都不是矩阵相乘，却利用矩阵相乘的法则来算	
数乘向量 向量乘数	向量乘数 不自然	数乘向量 自然
过渡矩阵 和 $f_i = Te_i$ 的 $M(T)$	过渡矩阵 $= M(T;(e_1,e_2))$	过渡矩阵 $= M(T;(e_1,e_2))$ 的转置 下页解释

$$[f_1, f_2] = [e_1, e_2]\underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \qquad \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}}_{B}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

观察 A 的 (i,j) 元 $=$ B 的 (j,i) 元

即 $A_{1,1} = B_{1,1}$

$A_{1,2} = \boxed{B_{2,1}}$

$\vdots$

这两个矩阵互称为对方的转置矩阵.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \xrightarrow{\text{把第一行转90度变成第一列}} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix}$$

$$[f_1, f_2] = [e_1, e_2]\underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \qquad \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}}_{B}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

观察 A 的 (i,j) 元 $=$ B 的 (j,i) 元

即 $A_{1,1} = B_{1,1}$

$A_{1,2} = \boxed{B_{2,1}}$

$\vdots$

这两矩阵互称为对方的转置矩阵。

$$\overset{A}{\begin{bmatrix} a_{11} & a_{12} \\ \boxed{a_{21} \quad a_{22}} \end{bmatrix}} \xrightarrow{\text{把第一行转90度变成第一列}} \overset{A^T}{\begin{bmatrix} a_{11} & \boxed{a_{21}} \\ a_{12} & \boxed{a_{22}} \end{bmatrix}}$$

解释矩阵的转置 (transposition).

$$[f_1, f_2] = [e_1, e_2]\underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \qquad \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}}_{B}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

观察 A 的 (i,j) 元 $=$ B 的 (j,i) 元

即 $A_{1,1} = B_{1,1}$

$A_{1,2} = \boxed{B_{2,1}}$

$\vdots$

这两个矩阵互称为对方的转置矩阵.

右上角放 T 是转置 (Transposition) 的首字母

A^T 表示 A 的转置

$$\overset{A}{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}} \xrightarrow{\text{把第一行转90度变成第一列}} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

坐标变换

(e_1, e_2) 是 V 的一个基，(f_1, f_2) 是 V 的另一个基，

那么一定存在一个线性映射 T，

使得 $f_1 = Te_1$

$f_2 = Te_2$

坐标变换

(e_1, e_2) 是 V 的一个基，(f_1, f_2) 是 V 的另一个基。

那么一定存在一个线性映射 T，

使得 $f_1 = Te_1$，$f_2 = Te_2$

$$\mathcal{M}(T; (e_1, e_2)) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

由前面的讨论知

$$[f_1, f_2] = [e_1, e_2] S,$$

其中 $S = \mathcal{M}(T; (e_1, e_2))$ 也是两个基之间的过渡矩阵。

如果有一个向量 $v = c_1 e_1 + c_2 e_2$，那么 v 在 (f_1, f_2) 的坐标是什么？即 Find $(d_1, d_2) \in F^2$，使得 $v = d_1 f_1 + d_2 f_2$.

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

$$e_1 = g_{11}f_1 + g_{21}f_2$$

$$e_2 = g_{12}f_1 + g_{22}f_2$$

$$M(T^{-1}; (f_1, f_2)) = \begin{bmatrix} g_{11} & g_{21} \\ g_{12} & g_{22} \end{bmatrix}$$

而 $$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$f_1 = a_{11}e_1 + a_{21}e_2$$

$$f_2 = a_{12}e_1 + a_{22}e_2$$

坐标变换

$$[f_1, f_2] = [e_1, e_2]S.$$

其中 $S = M(T;(e_1,e_2))$　$f_1 = Te_1$, $f_2 = Te_2$

T 可逆，$e_1 = T^{-1}f_1$, $e_2 = T^{-1}f_2$

$$M(T^{-1};(e_1,e_2)) = \left(M(T;(e_1,e_2))\right)^{-1} = S^{-1}$$

可是 $[e_1, e_2] = [f_1, f_2]\,Q$ → 是从 (f_1,f_2) 到 (e_1,e_2) 的过渡矩阵

这个过渡矩阵 $Q = S^{-1}$ 吗？

坐标变换

$$[f_1, f_2] = [e_1, e_2] S.$$

其中 $S = M(T; (e_1, e_2))$　　$f_1 = Te_1$, $f_2 = Te_2$

T 可逆，$e_1 = T^{-1} f_1$, $e_2 = T^{-1} f_2$

$$M(T^{-1}; (e_1, e_2)) = \left(M(T; (e_1, e_2))\right)^{-1} = S^{-1}$$

可是　$[e_1, e_2] = [f_1, f_2]\, Q$ 是从 (f_1, f_2) 到 (e_1, e_2) 的过渡矩阵

按照前面的做法 $Q = M(T^{-1}; (f_1, f_2))$

这个过渡矩阵 $Q = S^{-1}$ 吗？

坐标变换

$$[f_1, f_2] = [e_1, e_2] S.$$

其中 $S = M(T; (e_1, e_2))$ $f_1 = Te_1$, $f_2 = Te_2$

T 可逆, $e_1 = T^{-1} f_1$, $e_2 = T^{-1} f_2$

$$M(T^{-1}; (e_1, e_2)) = (M(T; (e_1, e_2)))^{-1} = S^{-1}$$

可是 $[e_1, e_2] = [f_1, f_2] Q$ 是从 (f_1, f_2) 到 (e_1, e_2) 的过渡矩阵

按照前面的做法 $Q = M(T^{-1}; (f_1, f_2))$ 这个过渡矩阵 $Q = S^{-1}$ 吗?

而 $S^{-1} = M(T^{-1}; (e_1, e_2))$

二者会相等吗? 这确实应该是个问题.

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$\boxed{\begin{aligned} e_1 &= q_{11} f_1 + q_{21} f_2 \\ e_2 &= q_{12} f_1 + q_{22} f_2 \end{aligned}}$$

$$M(T^{-1}; (f_1, f_2)) = \begin{bmatrix} q_{11} & q_{21} \\ q_{12} & q_{22} \end{bmatrix}$$

而 $$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\boxed{\begin{aligned} f_1 &= a_{11} e_1 + a_{21} e_2 \\ f_2 &= a_{12} e_1 + a_{22} e_2 \end{aligned}}$$

代入，得

$$\begin{aligned} e_1 &= q_{11}(a_{11} e_1 + a_{21} e_2) + q_{21}(a_{12} e_1 + a_{22} e_2) \\ &= (q_{11} a_{11} + q_{21} a_{12}) e_1 + (q_{11} a_{21} + q_{21} e_{22}) e_2 \\ &= e_1 \begin{bmatrix} a_{11} & a_{12} \end{bmatrix}\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} + e_2 \begin{bmatrix} a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} \end{aligned}$$

$$[e_1, e_2] = [f_1, f_2] \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$\boxed{\begin{aligned} e_1 &= q_{11} f_1 + q_{21} f_2 \\ e_2 &= q_{12} f_1 + q_{22} f_2 \end{aligned}}$$

$$M(T^{-1}; (f_1, f_2)) = \begin{bmatrix} q_{11} & q_{21} \\ q_{12} & q_{22} \end{bmatrix}$$

而 $$[f_1, f_2] = [e_1, e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

代入，得

$$\boxed{\begin{aligned} f_1 &= a_{11} e_1 + a_{21} e_2 \\ f_2 &= a_{12} e_1 + a_{22} e_2 \end{aligned}}$$

$$\begin{aligned} e_1 &= q_{11}(a_{11} e_1 + a_{21} e_2) + q_{21}(a_{12} e_1 + a_{22} e_2) \\ &= \underbrace{\boxed{(q_{11} a_{11} + q_{21} a_{12})}}_{=1} e_1 + \underbrace{\boxed{(q_{11} a_{21} + q_{21} e_{22})}}_{=0} e_2 \\ &= e_1 [a_{11} \ \ a_{12}] \begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} + e_2 [a_{21} \ \ a_{22}] \begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} \end{aligned}$$

即 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

上面的推导，每一步我们都是用完全严格法则运算.

即 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

上面的推导每一步我们都是用完全严格法则运算.

下面

$$[f_1, f_2] = [e_1, e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

这两个"矩阵"形式的写法是有点"毛病"，但是我们稍稍做了一点变通（如向量乘以数或数乘向量）就可以这样

即 $$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

上面的推导每一步我们都是用完全严格法则运算.

下面

$$[f_1, f_2] = [e_1, e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$[e_1, e_2] = [f_1, f_2] \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

这两个"矩阵"形式的写法是有点"毛病", 但是我们稍稍做了一点变通(如向量乘以数变成数乘向量)

就可以这样

$$[e_1, e_2] = [f_1, f_2] \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

$$= [e_1 \quad e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

上一页末的式子重抄一遍

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= [e_1 \ \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= [e_1 \ \ e_2]\begin{bmatrix} a_{11} q_{11} + a_{12} q_{21} & a_{11} q_{12} + a_{12} q_{22} \\ a_{21} q_{21} + a_{22} q_{21} & a_{21} q_{12} + a_{22} q_{22} \end{bmatrix}$$

上一页末的式子重抄一遍

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= [e_1 \ \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= \boxed{[e_1 \ \ e_2]}\begin{bmatrix} a_{11} q_{11} + a_{12} q_{21} & a_{11} q_{12} + a_{12} q_{22} \\ a_{21} q_{21} + a_{22} q_{21} & a_{21} q_{12} + a_{22} q_{22} \end{bmatrix}$$

$$e_1 = (a_{11} q_{11} + a_{12} q_{21}) e_1 + (a_{21} q_{21} + a_{22} q_{21}) e_2$$

上一页末的式子重抄一遍

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= [e_1 \ \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= \boxed{[e_1 \ \ e_2]}\begin{bmatrix} \boxed{\begin{matrix} a_{11}q_{11} + a_{12}q_{21} \\ a_{21}q_{21} + a_{22}q_{21} \end{matrix}} & \begin{matrix} a_{11}q_{12} + a_{12}q_{22} \\ a_{21}q_{12} + a_{22}q_{22} \end{matrix} \end{bmatrix}$$

$$e_1 = (a_{11}q_{11} + a_{12}q_{21})e_1 + (a_{21}q_{21} + a_{22}q_{21})e_2$$

从而，得 $a_{11}q_{11} + a_{12}q_{21} = [a_{11} \ \ a_{12}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 1$

$$a_{21}q_{21} + a_{22}q_{22} = [a_{21} \ \ a_{22}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 0$$

上一页末的式子重抄一遍

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= [e_1 \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= \boxed{[e_1 \ e_2]}\begin{bmatrix} a_{11}q_{11} + a_{12}q_{21} & a_{11}q_{12} + a_{12}q_{22} \\ a_{21}q_{21} + a_{22}q_{21} & a_{21}q_{12} + a_{22}q_{22} \end{bmatrix}$$

$$e_1 = (a_{11}q_{11} + a_{12}q_{21})e_1 + (a_{21}q_{21} + a_{22}q_{21})e_2$$

从而、可得 $a_{11}q_{11} + a_{12}q_{21} = [a_{11} \ a_{12}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 1$

$$a_{21}q_{21} + a_{22}q_{22} = [a_{21} \ a_{22}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 0$$

类似地，可得 $e_2 = [a_{11} \ a_{12}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} e_1 + [a_{21} \ a_{22}]\begin{bmatrix} q_{12} \\ q_{22} \end{bmatrix} e_2$

上一页末的式子重抄一遍

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= [e_1 \ \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= \boxed{[e_1 \ \ e_2]}\begin{bmatrix} a_{11}q_{11} + a_{12}q_{21} & a_{11}q_{12} + a_{12}q_{22} \\ a_{21}q_{21} + a_{22}q_{21} & a_{21}q_{12} + a_{22}q_{22} \end{bmatrix}$$

$$e_1 = (a_{11}q_{11} + a_{12}q_{21})\, e_1 + (a_{21}q_{21} + a_{22}q_{21})\, e_2$$

从而，可得 $a_{11}q_{11} + a_{12}q_{21} = [a_{11} \ \ a_{12}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 1$

$$a_{21}q_{21} + a_{22}q_{22} = [a_{21} \ \ a_{22}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 0$$

类似地，可得 $e_2 = e_1 [a_{11} \ \ a_{12}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} + e_2 [a_{21} \ \ a_{22}]\begin{bmatrix} q_{12} \\ q_{22} \end{bmatrix}$

上一页末的式子重抄一遍

$$[e_1, e_2] = [f_1, f_2]\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= [e_1 \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$= \boxed{[e_1 \ e_2]}\left[\boxed{\begin{matrix} a_{11}q_{11}+a_{12}q_{21} \\ q_{21}q_{21}+q_{22}q_{21} \end{matrix}} \ \begin{matrix} a_{11}q_{12}+a_{12}q_{22} \\ a_{21}q_{12}+a_{22}q_{22} \end{matrix}\right]$$

$$e_1 = (a_{11}q_{11}+a_{12}q_{21})e_1 + (a_{21}q_{21}+a_{22}q_{21})e_2$$

从而，可得 $a_{11}q_{11}+a_{12}q_{21} = [a_{11} \ a_{12}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 1$

$$a_{21}q_{21}+a_{22}q_{22} = [a_{21} \ a_{22}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix} = 0$$

类似地，可得 $e_2 = e_1\underbrace{[a_{11} \ a_{12}]\begin{bmatrix} q_{11} \\ q_{21} \end{bmatrix}}_{=0} + e_2\underbrace{[a_{21} \ a_{22}]\begin{bmatrix} q_{12} \\ q_{22} \end{bmatrix}}_{=1}$

上面的严格推导和"形式上用矩阵乘法"推导是一致的.
也就证明了用"矩阵乘法"推导的正确性.

更简单地写法

$$[f_1 \ \ f_2] = [e_1 \ \ e_2]\, S$$

这里的 S 是 Transitional matrix（过渡矩阵）

它和 $f_i = T e_i$ 的映射 T 在基 (e_1, e_2) 下的矩阵 $M(T;(e_1,e_2))$ 相等.

上面的严格推导和"形式上用矩阵乘法"推导是一致的.
也就证明用"矩阵乘法"推导的正确性.

更简单地写法

$$[f_1 \quad f_2] = [e_1 \quad e_2] S \qquad ①$$

这里的 S 是 Transitional matrix（过渡矩阵）

它和 $f_i = T e_i$ 的映射 T 在基 (e_1, e_2) 下的矩阵 $M(T; (e_1, e_2))$ 相等.

$$[e_1, e_2] = [f_1, f_2] Q \qquad ②$$

把②代入①, 得 $[e_1, e_2] = [e_1 \quad e_2] S Q$

从而得 $SQ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$. 所以 S 和 Q 互逆.

上面的严格推导和"形式上用矩阵乘法"推导是一致的。也就证明用"矩阵乘法"推导的正确性。

更简单地写法

$$[f_1 \quad f_2] = [e_1 \quad e_2]\, S \qquad (1)$$

这里的 S 是 Transitional matrix（过渡矩阵）

更简单地说，我们会明白 S 是什么基下的矩阵。

直解这样。(1) 式的两端同时右乘 S^{-1}，得

$$[f_1 \quad f_2]\, S^{-1} = [e_1 \quad e_2]\, S S^{-1}$$

即

$$[f_1 \quad f_2]\, S^{-1} = [e_1 \quad e_2].$$

上面的严格推导和"形式上用矩阵乘法"推导是一致的。
也就证明了用"矩阵乘法"推导的正确性。

更简单地写法

$$[f_1 \quad f_2] = [e_1 \quad e_2] S \qquad (1)$$

这里的 S 是 Transitional matrix（过渡矩阵）

(1)式的两端同时右乘 S^{-1}，得

$$[f_1 \quad f_2] S^{-1} = [e_1 \quad e_2] S S^{-1}$$

利用 $SS^{-1} = I$

即 $$[f_1 \quad f_2] S^{-1} = [e_1 \quad e_2]$$

这些都得被看作会法地自然地操作。

再回到坐标变换

求 $v = c_1 e_1 + c_2 e_2$ 在新的基 (f_1, f_2) 下的坐标，

这里 $[f_1 \ \ f_2] = [e_1 \ \ e_2] S$，$S$ 是从基 (e_1, e_2)

到基 (f_1, f_2) 下的过渡矩阵.

写成 $v = [e_1 \ \ e_2] \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

再回到坐标变换

求 $v = c_1 e_1 + c_2 e_2$ 在新的基 (f_1, f_2) 下的坐标，

这里 $[f_1 \ \ f_2] = [e_1 \ \ e_2] S$ ①，S是从基 (e_1, e_2)

到基 (f_1, f_2) 下的过渡矩阵.

写成 $v = [e_1 \ \ e_2] \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ ②

由①式，得 $[e_1 \ \ e_2] = [f_1 \ \ f_2] S^{-1}$ ③

再回到坐标变换

求 $v = c_1 e_1 + c_2 e_2$ 在新的基 (f_1, f_2) 下的坐标，

这里 $[f_1 \ f_2] = [e_1 \ e_2] S$ ①，S 是从基 (e_1, e_2)

到基 (f_1, f_2) 下的过渡矩阵。

写成 $v = [e_1 \ e_2] \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ ②

由①式，得 $[e_1 \ e_2] = [f_1 \ f_2] S^{-1}$ ③

把③代入②，得 $v = [f_1 \ f_2] S^{-1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

再回到坐标变换

求 $v = c_1 e_1 + c_2 e_2$ 在新的基 (f_1, f_2) 下的坐标，

这里 $[f_1 \quad f_2] = [e_1 \quad e_2] S$ ① ，S 是从基 (e_1, e_2) 到基 (f_1, f_2) 下的过渡矩阵.

写成 $v = [e_1 \quad e_2]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ ②

由①式，得 $[e_1 \quad e_2] = [f_1 \quad f_2] S^{-1}$ ③

把③代入②，得 $v = [f_1 \quad f_2] S^{-1}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

这就是 v 在基 (f_1, f_2) 下的坐标.

映射T的矩阵

(e_1, e_2) 是V的一个基. (f_1, f_2) 是V的另一个基,

$$[f_1 \quad f_2] = [e_1 \quad e_2]\, S$$

其中S 过渡矩阵.

$T \in \mathcal{L}(V)$ 是一个 linear map,

$$M(T; (e_1, e_2)) = \begin{array}{c c} & \begin{array}{cc} Te_1 & Te_2 \end{array} \\ \begin{array}{c} e_1 \\ e_2 \end{array} & \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{array} = A$$

映射T的矩阵

(e_1, e_2) 是V的一个基. (f_1, f_2) 是V的另一个基,

$$[f_1 \quad f_2] = [e_1 \quad e_2]\, S$$

其中 S 过渡矩阵

$T \in \mathcal{L}(V)$ 是一个 linear map.

$$\mathcal{M}(T, (e_1, e_2)) = \begin{array}{c|cc} & Te_1 & Te_2 \\ \hline e_1 & a_{11} & a_{12} \\ e_2 & a_{21} & a_{22} \end{array} = A$$

$$Te_1 = a_{11} e_1 + a_{21} e_2 = [e_1 \quad e_2]\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$$

$$Te_2 = a_{12} e_1 + a_{22} e_2 = [e_1 \quad e_2]\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

$$[Te_1 \quad Te_2] = [e_1 \quad e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

映射T的矩阵

(e_1, e_2) 是V的一个基. (f_1, f_2) 是V的另一个基,

$$[f_1 \quad f_2] = [e_1 \quad e_2] S$$

其中 S 过渡矩阵

$T \in \mathcal{L}(V)$ 是一个 linear map.

$$M(T, (e_1, e_2)) = \begin{array}{c|cc} & Te_1 & Te_2 \\ \hline e_1 & a_{11} & a_{12} \\ e_2 & a_{21} & a_{22} \end{array} = A$$

$$Te_1 = a_{11} e_1 + a_{21} e_2 = [e_1 \quad e_2] \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$$

$$Te_2 = a_{12} e_1 + a_{22} e_2 = [e_1 \quad e_2] \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

$$[Te_1 \quad Te_2] = [e_1 \quad e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$[Te_1 \ \ Te_2]$ 又可以写成 $T[e_1 \ \ e_2]$

同理. 上页页末两式子就可以写成

$$T[e_1 \ \ e_2] = [e_1 \ \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$[Te_1 \quad Te_2]$ 又可以写成 $T[e_1 \quad e_2]$

同理. 上页页末两式子就可以写成

$$T[e_1 \quad e_2] = [e_1 \quad e_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

即 $T[e_1 \quad e_2] = [e_1 \quad e_2] \underbrace{A}_{\mathcal{M}(T;(e_1,e_2))}$

$[Te_1 \quad Te_2]$ 又可以写成 $T[e_1 \quad e_2]$

同理. 上页最末两式子就可以写成

$$T[e_1 \quad e_2] = [e_1 \quad e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

即 $$T[e_1 \quad e_2] = [e_1 \quad e_2]\underbrace{A}_{\mathcal{M}(T;\,(e_1,\,e_2))}$$

$$\forall\ v = c_1e_1 + c_2e_2 = [e_1 \quad e_2]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = [e_1 \quad e_2]\,\vec{c}$$

$$Tv = ?$$

$[Te_1 \ \ Te_2]$ 又可以写成 $T[e_1 \ \ e_2]$

同理，上页变来而式子就可以写成

$$T[e_1 \ \ e_2] = [e_1 \ \ e_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

即 $\boxed{T[e_1 \ \ e_2] = [e_1 \ \ e_2]\underbrace{A}_{\mathcal{M}(T;(e_1,e_2))}}$

$\forall \ v = c_1e_1 + c_2e_2 = [e_1 \ \ e_2]\underbrace{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}} = [e_1 \ \ e_2]\underbrace{\vec{c}}$

$Tv = ?$

$Tv = T[e_1 \ \ e_2]\vec{c} = [e_1 \ \ e_2]A\vec{c}$

所以这和我们以前学的 $\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$ 是一致的

如何从一个基下的矩阵得到另一个基下的矩阵

设 $(e_1, \cdots, e_n)$ 和 $(f_1, \cdots, f_n)$ 是线性空间 V 的两个基, 我们下面用爱因斯坦求和约定.

基的变换

$$f_j = s_{k,j} e_k, \quad j = 1, \cdots, n.$$

$$e_j = t_{k,j} f_j, \quad j = 1, \cdots, n.$$

$$T[e_1,\cdots,e_n] = [e_1,\cdots,e_n]A$$

$$[f_1,\cdots,f_n] = [e_1,\cdots,e_n]S$$

求 B，使得

$$T[f_1,\cdots,f_n] = [f_1,\cdots,f_n]B$$

B 就是映射 $T\in\mathcal{L}(V)$ 在基 $(f_1,\cdots,f_n)$ 下的矩阵

$$T[e_1,\cdots,e_n] = [e_1,\cdots,e_n]A \qquad ①$$

$$[f_1,\cdots,f_n] = [e_1,\cdots,e_n]S \qquad ②$$

求 B，使得

$$T[f_1,\cdots,f_n] = [f_1,\cdots,f_n]B \qquad ③$$

B 就是映射 $T\in\mathcal{L}(V)$ 在基 $(f_1,\cdots,f_n)$ 下的矩阵

把②式代入③式的左端，得

$$T[f_1,\cdots,f_n] = T\big([e_1,\cdots,e_n]S\big)$$

$$T[e_1,\cdots,e_n]=[e_1,\cdots,e_n]A \quad ①$$

$$[f_1,\cdots,f_n]=[e_1,\cdots,e_n]S \quad ②$$

求 B，使得

$$T[f_1,\cdots,f_n]=[f_1,\cdots,f_n]B \quad ③$$

B 就是映射 $T\in\mathcal{L}(V)$ 在基 $(f_1,\cdots,f_n)$ 下的矩阵

把②式代入③式的左端，得

$$T[f_1,\cdots,f_n]=T\left([e_1,\cdots,e_n]S\right)$$

线性映射，可以去掉(　)

$$T[e_1,\cdots,e_n] = [e_1,\cdots,e_n]A \qquad ①$$

$$[f_1,\cdots,f_n] = [e_1,\cdots,e_n]S \qquad ②$$

求 B，使得

$$T[f_1,\cdots,f_n] = [f_1,\cdots,f_n]B \qquad ③$$

B 就是映射 $T\in\mathcal{L}(V)$ 在基 $(f_1,\cdots,f_n)$ 下的矩阵

把②式代入③式的左端，得

$$T[f_1 \cdots f_n] = T\big([e_1 \cdots e_n]S\big)$$

线性映射，可以去掉（ ）

$$= T[e_1 \cdots e_n]S$$

由 $T[e_1, \dots, e_n] = [e_1 \cdots e_n]A$，代入，可得

$$T[f_1 \cdots f_n] = [e_1 \cdots e_n]AS$$

用 $T[e_1, \dots, e_n] = [e_1 \cdots e_n]A$，代入，得

$$T[f_1 \cdots f_n] = [e_1 \cdots e_n]AS$$

再用 $[e_1, \cdots e_n] = [f_1 \cdots f_n]S^{-1}$ 代入，得

$$T[f_1 \cdots f_n] = [f_1 \cdots f_n]S^{-1}AS$$

故 $M(T, (f_1, \dots, f_n)) = S^{-1}AS$.

目标

我们知道 $\mathcal{M}(T,(e_1,\cdots,e_n)) = [A_{i,j}]$, 想求 $\mathcal{M}(T,(f_1,\cdots,f_n))$.

比较之后, 我们就发现

新旧的基下的矩阵的关系

$$\mathcal{M}(T,(f_1,\cdots,f_n)) = \underbrace{S^{-1}}_{[e_1,\cdots,e_n]=[f_1,\cdots,f_n]S^{-1}} \mathcal{M}(T,(e_1,\cdots,e_n))\, S$$

矩阵的逆

关于矩阵的逆, 我们有下面的补充, 需要大家复习一下算子的逆. 可以想象, 一个矩阵的逆, 就是它对应的算子的逆的矩阵. 设 $A \in F^{n\times n}$ 是一个 $n \times n$ 的矩阵, 如果存在 $B \in F^{n\times n}$ 满足

$$AB = I, \quad BA = I,$$

其中 $I \in F^{n\times n}$ 是单位矩阵, 我们称 B 是的 A 的逆. 可以证明一个矩阵如果逆存在, 则这个逆是唯一的. 因此可以把矩阵的逆记作 A^{-1}.

$$(AB)_{i,j} = A_{i,k}B_{k,j}$$

使用了爱因斯坦求和约定 $A_{i,k}B_{k,j} = \sum_{k=1}^{n} A_{i,k}B_{k,j}$.

" 正交性"

我们可以看到 A 和 A^{-1} 满足

$$A_{i,k}(A^{-1})_{k,j} = \begin{cases} 1, & \text{if i = j} \\ 0, & \text{else} \end{cases}$$

再看 Gram-Schmidt 正交化 I

设 (v_1, v_2, v_3) 是三维线性空间的一个基, Gram-Schmidt procedure 给出了一个正交基

$$
\begin{aligned}
e_1 &= \frac{v_1}{\|v_1\|} \\
e_2 &= \frac{v_2 - \langle v_2, e_1\rangle e_1}{\|v_2 - \langle v_2, e_1\rangle e_1\|} \\
e_3 &= \frac{v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2}{\|v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2\|}
\end{aligned}
$$

从基 (v_1, v_2, v_3) 到基 (e_1, e_2, e_3) 的变换的矩阵是

$$
\begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} S_{1,1} & S_{1,2} & S_{1,3} \\ 0 & S_{2,2} & S_{2,3} \\ 0 & 0 & S_{2,3} \end{bmatrix}
$$

观察上面的矩阵

再看 Gram-Schmidt 正交化 II

观察到 Gram-Schmidt 正交化使用的变换矩阵是一个三角矩阵. 从这个矩阵, 我们应该可以到 e_1 仅仅使用 v_1 来线性表出, e_2 仅仅使用 v_1, v_2 来线性表出, e_3 仅仅使用 v_1, v_2, v_3 来线性表出,

那我们把用 (v_1, v_2, v_3) 来表示 (e_1, e_2, e_3)

$$\begin{aligned} v_1 &= \|v_1\| e_1 \\ v_2 &= \|v_2 - \langle v_2, e_1 \rangle e_1\| e_2 + \langle v_2, e_1 \rangle e_1 \\ v_3 &= \|v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2\| e_3 + \langle v_3, e_1 \rangle e_1 + \langle v_3, e_2 \rangle e_2 \end{aligned}$$

从基 (e_1, e_2, e_3) 到基 (v_1, v_2, v_3) 的变换的矩阵是

$$\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} \begin{bmatrix} \left(S^{-1}\right)_{1,1} & \left(S^{-1}\right)_{1,2} & \left(S^{-1}\right)_{1,3} \\ 0 & \left(S^{-1}\right)_{2,2} & \left(S^{-1}\right)_{2,3} \\ 0 & 0 & \left(S^{-1}\right)_{2,3} \end{bmatrix}$$

观察上面的矩阵

再看 Gram-Schmidt 正交化 III

从这个矩阵, 我们应该可以到 v_1 仅仅使用 e_1 来线性表出, v_2 仅仅使用 e_1, e_2 来线性表出, v_3 仅仅使用 e_1, e_2, e_3 来线性表出,

观察: 上三角矩阵的逆一定是上三角矩阵

另外, 我们还看到一个上三角矩阵如果是可逆的, 那么它的逆一定是上三角矩阵.

两个上三角矩阵的乘积还是上三角矩阵

设 $A = (A_{i,j}), B = (B_{i,j})$, 满足

$$A_{i,j} = 0, \text{ if j>i}$$
$$B_{i,j} = 0, \text{ if j>i}$$

那么 $(AB)_{i,j} = 0, \text{ if j >i }.$

Proof.

再看 Gram-Schmidt 正交化 IV

$$(AB)_{i,j} = A_{i,k}B_{k,j}$$

我们检查一下是不是当 $j > i$ 的时候, 分两种情况,
case 1: $k < j$, 那么一定有 $B_{k,j} = 0$,
case 2: $k \geq j$, 那么一定有 $k > i$(recall that $j > i$), 于是一定有 $A_{i,k} = 0$
总之,$A_{i,k}B_{k,j}$ 在 $j > i$ 的时候总是为 0. 于是我们就证明了当 $j > i$ 时,$(AB)_{i,j} = 0$ □

推论 6.27 的证明

6.27 推论：设 V 是一个有限维非零内积空间，$T \in \mathcal{L}(V)$. 如果 T 关于 V 的某个基(basis)具有上三角矩阵，那么 T 关于 V 的某个规范正交基 (orthogonal basis) 也具有上三角矩阵.

Proof.

算子 $T \in \mathcal{L}(V)$ 关于 $(v_1, \cdots, v_n)$ 具有上三角矩阵, Gram-Schmidt 正交化得到一个规范正交基 $(e_1, \cdots, e_n)$ 满足

$$\begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} S$$

其中 S 是上三角可逆矩阵, 而且 S^{-1} 也是上三角矩阵.

$$\mathcal{M}\left(T, (e_1, \cdots, e_n)\right) = S^{-1} \mathcal{M}\left(T, (v_1, \cdots, v_n)\right) S$$

再利用上三角矩阵的乘积还是上三角矩阵, 我们从 T 关于某个基具有上三角矩阵推出了存在一个规范正交基, 使得 T 在该规范正交基下的矩阵是上三角矩阵.

Gram-Schmidt 正交化

算子的规范正交上三角化

坐标变换

正交投影和极小化问题

作业

正交投影和极小化问题

§6.4 正交投影与极小化问题

如果 U 是 V 的子集, 那么 U 的**正交补** (**orthogonal complement**), 记为 $U^{\perp}$, 是由 V 中与 U 的每个向量都正交的那些向量组成的集合:

$$U^{\perp} = \{\boldsymbol{v} \in V : \langle \boldsymbol{v}, \boldsymbol{u} \rangle = 0, \boldsymbol{u} \in U\}.$$

你应该验证, $U^{\perp}$ 总是 V 的子空间, $V^{\perp} = \{\mathbf{0}\}$, 并且 $\{\mathbf{0}\}^{\perp} = V$. 注意到, 若 $U_1 \subset U_2$, 则 $U_1^{\perp} \supset U_2^{\perp}$.

回想一下，如果 U_1, U_2 都是 V 的子空间，并且 V 中每个元素都可以唯一地写成 U_1 中的一个向量与 U_2 中的一个向量的和，那么 V 是 U_1 和 U_2 的直和 (记为 $V = U_1 \oplus U_2$). 下一个定理表明，由内积空间的每个子空间都可以得到整个空间的一个自然的直和分解.

6.29 定理：如果 U 是 V 的子空间，那么

$$V = U \oplus U^{\perp}.$$

证明：设 U 是 V 的子空间. 首先证明

6.31

$$\boldsymbol{v}=\underbrace{\langle\boldsymbol{v},\boldsymbol{e}_1\rangle\boldsymbol{e}_1+\cdots+\langle\boldsymbol{v},\boldsymbol{e}_m\rangle\boldsymbol{e}_m}_{\boldsymbol{u}}+\underbrace{\boldsymbol{v}-\langle\boldsymbol{v},\boldsymbol{e}_1\rangle\boldsymbol{e}_1-\cdots-\langle\boldsymbol{v},\boldsymbol{e}_m\rangle\boldsymbol{e}_m}_{\boldsymbol{w}}.$$

显然 $\boldsymbol{u}\in U$. 因为 $(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m)$ 是一个规范正交组, 所以对每个 j 都有

$$\begin{aligned}\langle\boldsymbol{w},\boldsymbol{e}_j\rangle&=\langle\boldsymbol{v},\boldsymbol{e}_j\rangle-\langle\boldsymbol{v},\boldsymbol{e}_j\rangle\\&=0.\end{aligned}$$

因此, $\boldsymbol{w}$ 正交于 $\operatorname{span}(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m)$ 中的每个向量. 即 $\boldsymbol{w}\in U^{\perp}$. 于是 $\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{w}$, 其中 $\boldsymbol{u}\in U$, $\boldsymbol{w}\in U^{\perp}$, 这就证明了 6.30.

若 $\boldsymbol{v} \in U \cap U^{\perp}$, 则 $\boldsymbol{v}$ (包含于 U) 正交于 U 中每个向量 (其中包括 $\boldsymbol{v}$ 本身), 从而 $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$, 即 $\boldsymbol{v} = \mathbf{0}$. 因此

6.32
$$U \cap U^{\perp} = \{0\}.$$

由 6.30 和 6.32 可得 $V = U \oplus U^{\perp}$ (参见 1.9). ■

下一个推论是定理 6.29 的重要结果.

6.33 推论: 如果 U 是 V 的子空间, 那么

$$U = (U^{\perp})^{\perp}.$$

证明：设 U 是 V 的子空间。首先证明

$$U \subset (U^{\perp})^{\perp}.$$

为此，设 $u \in U$，则对每个 $v \in U^{\perp}$ 都有 $\langle u, v\rangle = 0$

（根据 $U^{\perp}$ 的定义，$v \in U^{\perp}$，那么 v 与 U 中的每一个元素都正交）

证明：设 U 是 V 的子空间。首先证明

$$\boxed{U \subset (U^\perp)^\perp.}$$

为此，设 $u \in U$，则对每个 $v \in U^\perp$ 都有 $\langle u, v\rangle = 0$

（根据 $U^\perp$ 的定义，$v \in U^\perp$，那么 v 与 U 中的每一个元素都正交）

因为 u 正交于 $U^\perp$ 中的每个向量，所以 $u \in (U^\perp)^\perp$，这就证明了

证明：设 U 是 V 的子空间。首先证明

$$\boxed{U \subset (U^\perp)^\perp.}$$

为此，设 $u \in U$，则对每个 $v \in U^\perp$ 都有 $\langle u, v\rangle = 0$

（根据 $U^\perp$ 的定义，$v \in U^\perp$，那么 v 与 U 中的每一个元素都正交）

因为 u 正交于 $U^\perp$ 中的每个向量，所以 $u \in (U^\perp)^\perp$，这就证明了

下面再证明 $(U^\perp)^\perp \subset U$.

要证明 $(U^\perp)^\perp \subset U$。 $\forall v \in (U^\perp)^\perp$，由子空间的正交补空间的定义，$(U^\perp)^\perp = \{v \in V : \langle v, u \rangle = 0,\ \text{for all } u \in U^\perp\}$

也就是说 $(U^\perp)^\perp$ 由所有的与 $U^\perp$ 中的所有向量都正交的向量组成。

要证明 $(U^\perp)^\perp \subset U$。 $\forall v \in (U^\perp)^\perp$，由子空间的正交补空间的定义

$$(U^\perp)^\perp = \{v \in V : \langle v, u\rangle = 0, \text{ for all } u \in U^\perp\}$$

也就是说 $(U^\perp)^\perp$ 由所有的与 $U^\perp$ 中的所有向量都正交的向量组成。

但是否 $\forall v \in (U^\perp)^\perp$ 就有 $v \in U$ 呢？ we don't know yet. 然而我们总可以将 v 作下面的正交分解

$$v = u + w$$

其中 $u \in U$，$w \in U^\perp$

我们如果能证明 $w = 0$，那么就完成了 $v \in U$ 的证明。

要证明 $(U^\perp)^\perp \subset U$。 $\forall v \in (U^\perp)^\perp$，由子空间的正交补空间的定义，

$$(U^\perp)^\perp = \{v \in V : \langle v, u\rangle = 0, \text{ for all } u \in U^\perp\}$$

也就是说 $(U^\perp)^\perp$ 由所有的与 $U^\perp$ 中的所有向量都正交的向量组成。

但是否 $\forall u \in (U^\perp)^\perp$ 就有 $v \in U$ 呢？ we don't know yet. 然而我们总可以将 v 作下面的正交分解

$$v = u + w$$

其中 $u \in U$，$w \in U^\perp$

我们如果能证明 $w=0$，那么就完成了 $v \in U$ 的证明。

利用 $(U^\perp)^\perp$ 的定义知，$v \in (U^\perp)^\perp$，那么 v 与 $U^\perp$ 中的任何元素都正交，而 $w \in U^\perp$，自然，也有 $\langle v, w\rangle = 0$。

要证明 $(U^\perp)^\perp \subset U$. $\forall v \in (U^\perp)^\perp$, 由子空间的正交补空间的定义 $(U^\perp)^\perp = \{v \in V : \langle v, u\rangle = 0, \text{ for all } u \in U^\perp\}$

也就是说 $(U^\perp)^\perp$ 由所有的与 $U^\perp$ 中的所有向量都正交的向量组成。

但是否 $\forall u \in (U^\perp)^\perp$ 就有 $v \in U$ 呢? we don't know yet. 然而我们总可以将 v 作下面的正交分解

$$v = u + w$$

其中 $u \in U$, $w \in U^\perp$

我们如果能证明 $w = 0$, 那么就完成了 $v \in U$ 的证明。

①

代入

利用 $(U^\perp)^\perp$ 的定义知, $v \in (U^\perp)^\perp$, 那么 v 与 $U^\perp$ 中的任何元素都正交, 而 $w \in U^\perp$, 自然, 也有 $\langle v, w\rangle = 0$。 而 $\langle v, w\rangle = \langle u+w, w\rangle$

要证明 $(U^\perp)^\perp \subset U$。 $\forall v \in (U^\perp)^\perp$，由子空间的正交补空间的定义

$$(U^\perp)^\perp = \{ v \in V : \langle v, u \rangle = 0, \text{ for all } u \in U^\perp \}$$

也就是说 $(U^\perp)^\perp$ 由所有的与 $U^\perp$ 中的所有向量都正交的向量组成。

但是否 $\forall v \in (U^\perp)^\perp$ 就有 $v \in U$ 呢？ we don't know yet. 然而我们总可以将 v 作下面的正交分解

$$v = u + w$$

其中 $u \in U$，$w \in U^\perp$

我们如果能证明 $w = 0$，那么就完成了 $v \in U$ 的证明。

利用 $(U^\perp)^\perp$ 的定义知，$v \in (U^\perp)^\perp$，那么 v 与 $U^\perp$ 中的任何元素都正交，而 $w \in U^\perp$，自然，也有 $\langle v, w \rangle = 0$。 ① 代入 而 $\langle v, w \rangle = \langle u+w, w \rangle =$ 下行

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

$u \in U$
$w \in U^\perp$ 所
总得 $\langle u, w \rangle = 0$

要证明 $(U^{\perp})^{\perp} \subset U$. $\forall v \in (U^{\perp})^{\perp}$, 由子空间的正交补空间的定义. $(U^{\perp})^{\perp} = \{ v \in V : \langle v, u \rangle = 0, \text{ for all } u \in U^{\perp} \}$

也就是说 $(U^{\perp})^{\perp}$ 由所有的与 $U^{\perp}$ 中的所有向量都正交的向量组成。

但是否 $\forall u \in (U^{\perp})^{\perp}$ 就有 $v \in U$ 呢? we don't know yet. 然而我们总可以将 v 作下面的正交分解

① 代入

$$v = u + w$$

其中 $u \in U$, $w \in U^{\perp}$

我们如果能证明 $w = 0$, 那么就完成了 $v \in U$ 的证明。

利用 $(U^{\perp})^{\perp}$ 的定义知. $v \in (U^{\perp})^{\perp}$, 那么 v 与 $U^{\perp}$ 中的任何元素都正交, 而 $w \in U^{\perp}$, 自然, 也有 $\langle v, w \rangle = 0$. ② 而 $\langle v, w \rangle = \langle u + w, w \rangle =$ 下行

$\langle u, w \rangle + \langle w, w \rangle = \langle w, w \rangle$ ③ 由 ② ③ 得 $w = 0$.

$u \in U$
$w \in U^{\perp}$
证得 $\langle u, w \rangle = 0$

要证明 $(U^\perp)^\perp \subset U$. $\forall v \in (U^\perp)^\perp$, 由子空间的正交补空间的定义. $(U^\perp)^\perp = \{ v \in V : \langle v, u \rangle = 0, \text{ for all } u \in U^\perp \}$

也就是说 $(U^\perp)^\perp$ 由所有的与 $U^\perp$ 中的所有向量都正交的向量组成。

但是否 $\forall v \in (U^\perp)^\perp$ 就有 $v \in U$ 呢? we don't know yet. 然而我们总可以将 v 作下面的正交分解

$$\boxed{v = u + w}$$

① 代入

其中 $u \in U$, $w \in U^\perp$

我们如果能证明 $w = 0$, 那么就完成了 $v \in U$ 的证明.

利用 $(U^\perp)^\perp$ 的定义知, $v \in (U^\perp)^\perp$, 那么 v 与 $U^\perp$ 中的任何元素都正交, 而 $w \in U^\perp$, 自然, 也有 $\langle v, w \rangle = 0$. ② 而 $\langle v, w \rangle = \langle u + w, w \rangle =$ 下行

$\langle u, w \rangle + \langle w, w \rangle = \langle w, w \rangle$ ③ 由 ② ③, 得 $\boxed{w = 0}$. 于是 $v = u \in U$. □

$u \in U$, $w \in U^\perp$ 所以

⇒ 记得 $\langle u, w \rangle = 0$

设 U 是 V 的子空间. 6.29 所给出的分解 $V=U\oplus U^{\perp}$ 表明, 每个向量 $\boldsymbol{v}\in V$ 都可以唯一地写成

$$\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{w},$$

其中 $\boldsymbol{u}\in U$, $\boldsymbol{w}\in U^{\perp}$. 利用此分解来定义 V 上的算子 P_U, 称为 V 到 U 上的**正交投影** (**orthogonal projection**): 对 $\boldsymbol{v}\in V$, 定义 $P_U\boldsymbol{v}$ 为上面分解中的向量 $\boldsymbol{u}$. 采用所引进的记号则有 $P_U=P_{U,U^{\perp}}$. 你应该验证, $P_U\in\mathcal{L}(V)$, 并且具有以下性质:

- range $P_U=U$;
- null $P_U=U^{\perp}$;
- 对每个 $\boldsymbol{v}\in V$ 都有 $\boldsymbol{v}-P_U\boldsymbol{v}\in U^{\perp}$;
- ${P_U}^2=P_U$;
- 对每个 $\boldsymbol{v}\in V$ 都有 $\|P_U\boldsymbol{v}\|\leqslant\|\boldsymbol{v}\|$.

进一步, 由 6.29 的证明中所使用的分解 6.31 可知, 如果 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m)$ 是 U 的规范正交基, 那么对每个 $\boldsymbol{v} \in V$ 都有

6.35
$$P_U \boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 + \cdots + \langle \boldsymbol{v}, \boldsymbol{e}_m \rangle \boldsymbol{e}_m.$$

经常会遇到这样的问题: 给定 V 的子空间 U 和点 $\boldsymbol{v} \in V$, 求点 $\boldsymbol{u} \in U$ 使得 $\|\boldsymbol{v} - \boldsymbol{u}\|$ 最小. 下一个命题表明, 通过取 $\boldsymbol{u} = P_U \boldsymbol{v}$ 就可以解决这个极小化问题.

6.36 命题: 设 U 是 V 的子空间, 并且 $\boldsymbol{v} \in V$, 则

$$\|\boldsymbol{v} - P_U \boldsymbol{v}\| \leqslant \|\boldsymbol{v} - \boldsymbol{u}\|, \quad \boldsymbol{u} \in U.$$

进一步, 若 $\boldsymbol{u} \in U$ 使得上面的不等式是等式, 则 $\boldsymbol{u} = P_U \boldsymbol{v}$.

证明：设 $\boldsymbol{u} \in U$，则

$$\textbf{6.37} \qquad \|\boldsymbol{v} - P_U\boldsymbol{v}\|^2 \leqslant \|\boldsymbol{v} - P_U\boldsymbol{v}\|^2 + \|P_U\boldsymbol{v} - \boldsymbol{u}\|^2$$

$$\textbf{6.38} \qquad = \|(\boldsymbol{v} - P_U\boldsymbol{v}) + (P_U\boldsymbol{v} - \boldsymbol{u})\|^2$$

$$= \|\boldsymbol{v} - \boldsymbol{u}\|^2,$$

其中 6.38 由勾股定理 (6.3) 得到，之所以能用勾股定理是因为 $\boldsymbol{v} - P_U\boldsymbol{v} \in U^\perp$ 并且 $P_U\boldsymbol{v} - \boldsymbol{u} \in U$. 再取平方根即得要证的不等式.

不等式中的等号成立当且仅当 6.37 是等式，当且仅当 $\|P_U\boldsymbol{v} - \boldsymbol{u}\| = 0$，当且仅当 $\boldsymbol{u} = P_U\boldsymbol{v}$. ■

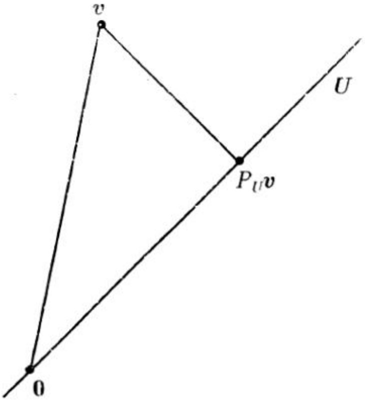
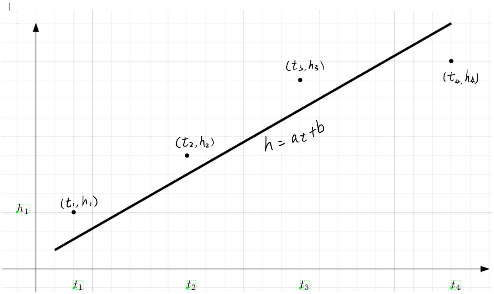


图 6-5　$P_U \boldsymbol{v}$ 是 U 中离 $\boldsymbol{v}$ 最近的点

数据拟合–最小二乘法

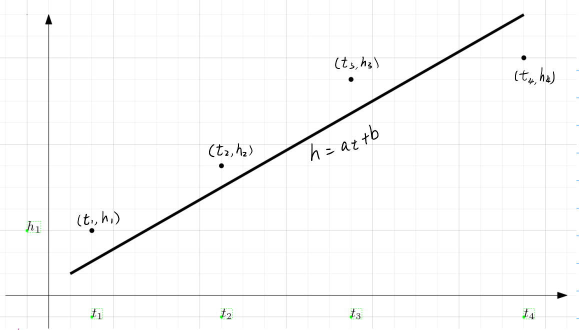


举例 1

测量得到一组数据
$(t_1, h_1), (t_2, h_2), (t_3, h_3), (t_4, h_4)$,
寻找一条直线
$h(t; a, b) = at + b, a, b \in \mathbb{R}$
使得下面的误差的函数

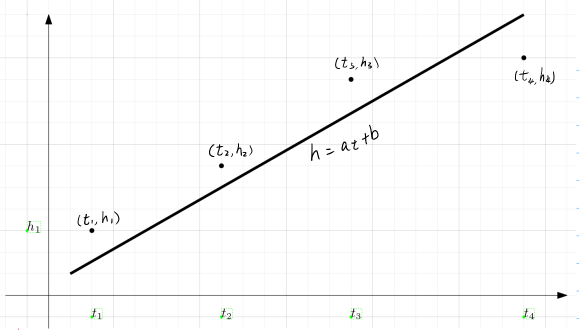
$$E(a, b) = \sum_{i=1}^{4} \|h(t_i; a, b) - h_i\|^2$$

最小.



求 $(a,b) \in \mathbb{R}^2$，使得 $\sum_{i=1}^{4} |h_i - (at_i+b)|^2$ 最小

定义一个映射 $T : \mathbb{R}^2 \to \mathbb{R}^4$ by $T(a,b) = (at_1+b, at_2+b, at_3+b, at_4+b)$

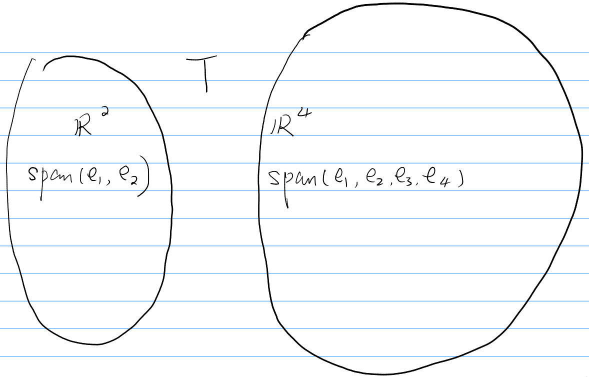


$\boxed{\text{求 } (a,b)\in\mathbb{R}^2\text{，使得 } \sum_{i=1}^{4} |h_i - (at_i+b)|^2 \text{ 最小}}$

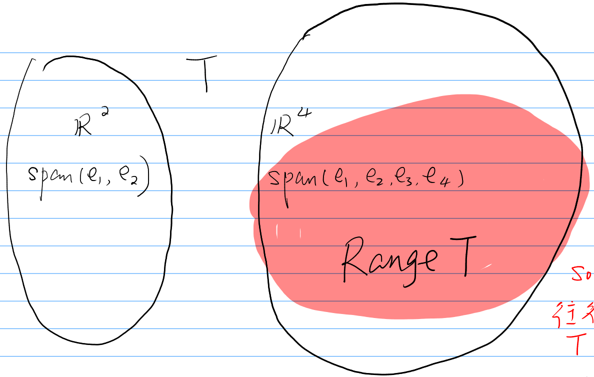
定义一个映射 $T:\mathbb{R}^2\to\mathbb{R}^4$ by $T(a,b) = (at_1+b, at_2+b, at_3+b, at_4+b)$

转化成：求 $(a,b)\in\mathbb{R}^2$，使得 $\|\vec{h} - T(a,b)\|^2$ 最小

其中 $\vec{h} = (h_1, h_2, h_3, h_4)$



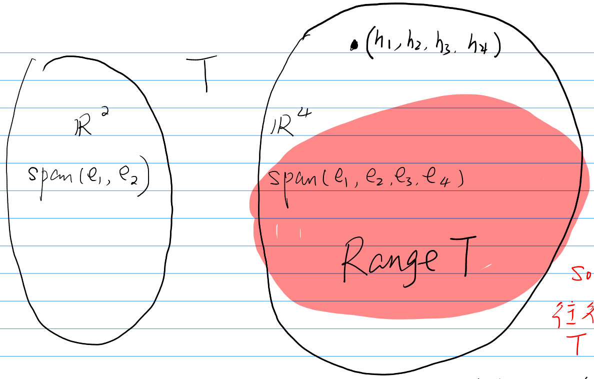
$$T(ae_1+be_2) = (at_1+b)e_1 + (at_2+b)e_2 + (at_3+b)e_3 + (at_4+b)e_4$$



T不是满射

So, (h_1, h_2, h_3, h_4) 往往不在Range T中

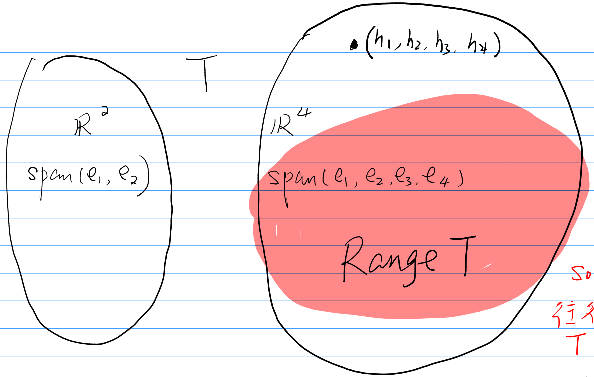
$$T(ae_1+be_2) = (at_1+b)e_1 + (at_2+b)e_2 + (at_3+b)e_3 + (at_4+b)e_4$$



T不是满射

So, (h_1, h_2, h_3, h_4) 往往不在Range T中

$$T(ae_1 + be_2) = (at_1 + b)e_1 + (at_2 + b)e_2 + (at_3 + b)e_3 + (at_4 + b)e_4$$

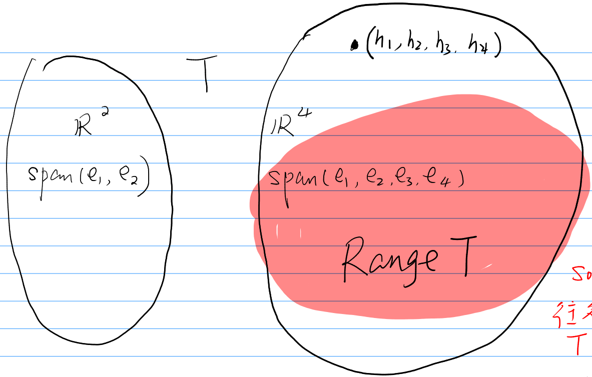


T不是满射

So, (h_1, h_2, h_3, h_4) 往往不在Range T中

$$T(ae_1+be_2) = (at_1+b)e_1+(at_2+b)e_2+(at_3+b)e_3+(at_4+b)e_4$$

问题：在 Range T 中与 (h_1, h_2, h_3, h_4) 最靠近的点是谁？

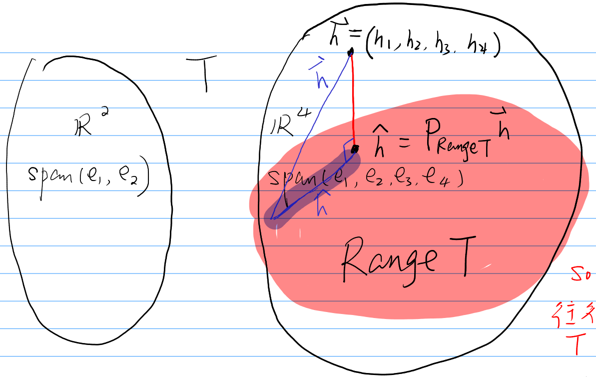


T不是满射
So, (h_1, h_2, h_3, h_4) 往往不在Range T中

$$T(ae_1+be_2) = (at_1+b)e_1 + (at_2+b)e_2 + (at_3+b)e_3 + (at_4+b)e_4$$

问题：在 Range T 中与 $\underbrace{(h_1, h_2, h_3, h_4)}_{\text{记作 } \vec{h}}$ 最靠近的点是谁？

命题6.36告诉我们 Range T 中与 $\vec{h}$ 最近的点是 $\vec{h}$ 在 Range T 中的正交投影 $P_{\text{Range } T}\vec{h}$



T不是满射

So, (h_1, h_2, h_3, h_4) 往往不在Range T中

$$T(ae_1+be_2) = (at_1+b)e_1 + (at_2+b)e_2 + (at_3+b)e_3 + (at_4+b)e_4$$

问题：在 Range T 中与 (h_1, h_2, h_3, h_4) 最靠近的点是谁?

记作 $\vec{h}$

命题6.36告诉我们 Range T 中与 $\vec{h}$ 最近的点是 $\vec{h}$ 在 Range T 中的正交投影 $P_{\mathrm{Range}\,T}\,\vec{h}$，记作 $\hat{h}$

$\vec{h} = (h_1, h_2, h_3, h_4) \in \mathbb{R}^4$

$T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^4)$ defined by $T(a,b) = (at_1+b, \dots, at_4+b)$

Step 1. 求 $\hat{h} = P_{\text{Range}\,T}\, \vec{h}$

Step 2: 求 $(a,b) = T^{-1}\hat{h}$

$\vec{h} = (h_1, h_2, h_3, h_4) \in \mathbb{R}^4$

$T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^4)$ defined by $T(a,b) = (at_1+b, \dots, at_4+b)$

Step 1. 求 $\hat{h} = P_{\text{Range}\,T}\,\vec{h}$

Step 2: 求 $(a,b) = \boxed{T^{-1}}\,\hat{h}$

T是否可逆呢？

T可能不是单射，但 $\hat{h} \in \text{Range}\,T$，所以一定是存在 $(a,b) \in \mathbb{R}^2$，使得 $T(a,b) = \hat{h}$ 而。

$\vec{h} = (h_1, h_2, h_3, h_4) \in \mathbb{R}^4$

$T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^4)$ defined by $T(a,b) = (at_1+b, \cdots, at_4+b)$

Step 1. 求 $\hat{h} = P_{\text{Range}\,T}\,\vec{h}$

Step 2: 求 $(a,b) = \boxed{T^{-1}}\,\hat{h}$

T 是否可逆呢?

由 Rank-Nullity theorem ($2 = \dim \operatorname{Null} T + \dim \operatorname{Range} T$)

知 如果 $\operatorname{Null} T = \{0\}$. 那么 T 可逆.

$\vec{h} = (h_1, h_2, h_3, h_4) \in \mathbb{R}^4$

$T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^4)$ defined by $T(a,b) = (at_1+b, \cdots, at_4+b)$

Step 1. 求 $\hat{h} = P_{\text{Range}\,T}\, \vec{h}$

Step 2: 求 $(a,b) = \boxed{T^{-1}}\, \hat{h}$

T是否可逆呢？

由 Rank-Nullity theorem ($2 = \dim \text{Null}\, T + \dim \text{Range}\, T$)

知 如果 $\text{Null}\, T = \{0\}$. 那么 T 可逆.

T不是单射的一种情况: $t_1 = t_2 = t_3 = t_4 = 0$,

此时. $T(a,b) = (b, b, b, b)$ Range T 是 $\mathbb{R}^4$ 中的一维子空间,

故由 Rank-nullity 定理 知 Null T 是 $\mathbb{R}^2$ 中的一维子空间.

Gram-Schmidt 正交化

算子的规范正交上三角化

坐标变换

正交投影和极小化问题

作业

作业

作业

8. 向量空间 U 上的范数是一个函数 $\|\ \|: U \to [0, \infty)$, 并且具有以下性质: $\|\boldsymbol{u}\| = 0$ 当且仅当 $\boldsymbol{u} = 0$; 对所有 $\alpha \in \mathbf{F}$ 和 $\boldsymbol{u} \in U$ 都有 $\|\alpha \boldsymbol{u}\| = |\alpha| \|\boldsymbol{u}\|$; 对所有 $\boldsymbol{u}, \boldsymbol{v} \in U$ 都有 $\|\boldsymbol{u} + \boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$. 证明: 满足平行四边形等式的范数都来自内积 (即证明: 如果 $\|\ \|$ 是 U 上的一个满足平行四边形等式的范数, 则有 U 上的内积 $\langle\ \rangle$, 使得对所有 $\boldsymbol{u} \in U$ 都有 $\|\boldsymbol{u}\| = \langle \boldsymbol{u}, \boldsymbol{u} \rangle^{1/2}$).

9. 设 n 是正整数. 证明

$$\left(\frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \cdots, \frac{\sin nx}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \cdots, \frac{\cos nx}{\sqrt{\pi}}\right)$$

是 $C[-\pi, \pi]$ 中的规范正交向量组, 其中 $C[-\pi, \pi]$ 是 $[-\pi, \pi]$ 上的连续实值函数所组成的向量空间, 其上的内积为

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \int_{-\pi}^{\pi} \boldsymbol{f}(x)\boldsymbol{g}(x)\mathrm{d}x.$$

10. 在 $\mathcal{P}_2(\mathbf{R})$ 上考虑内积

$$\langle \boldsymbol{p}, \boldsymbol{q} \rangle = \int_0^1 \boldsymbol{p}(x)\boldsymbol{q}(x)\mathrm{d}x.$$

对基 $(1, x, x^2)$ 应用格拉姆–施密特过程求 $\mathcal{P}_2(\mathbf{R})$ 的一个规范正交基.

11. 对非线性无关的向量组应用格拉姆–施密特过程结果会怎样?

12. 设 V 是实内积空间, 并且 $(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m)$ 是 V 中的线性无关向量组. 证明 V 中恰好有 2^m 个规范正交向量组 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m)$ 使得

$$\operatorname{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_j) = \operatorname{span}(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_j), \quad j \in \{1, \cdots, m\}.$$

13. 设 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m)$ 是 V 中的规范正交向量组, 并且 $\boldsymbol{v} \in V$. 证明

$$\|\boldsymbol{v}\|^2 = |\langle \boldsymbol{v}, \boldsymbol{e}_1 \rangle|^2 + \cdots + |\langle \boldsymbol{v}, \boldsymbol{e}_m \rangle|^2$$

当且仅当 $\boldsymbol{v} \in \operatorname{span}(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m)$.

14. 求 $\mathcal{P}_2(\mathbf{R})$ 的一个规范正交基 (内积如习题 10), 使得 $\mathcal{P}_2(\mathbf{R})$ 上的微分算子 (将 $\boldsymbol{p}$ 映成 $\boldsymbol{p}'$ 的算子) 关于此基具有上三角矩阵.

15. 设 U 是 V 的子空间. 证明

$$\dim U^{\perp} = \dim V - \dim U.$$

16. 设 U 是 V 的子空间. 证明 $U^{\perp} = \{\mathbf{0}\}$ 当且仅当 $U = V$.

17. 证明：如果 $P \in \mathcal{L}(V)$, 使得 $P^2 = P$ 并且 P 的零空间中的每个向量都正交于 P 的值域中的向量, 那么 P 是正交投影.

18. 证明: 如果 $P \in \mathcal{L}(V)$, 使得 $P^2 = P$ 并且

$$\|P\boldsymbol{v}\| \leqslant \|\boldsymbol{v}\|, \quad \boldsymbol{v} \in V,$$

那么 P 是正交投影.

19. 设 $T \in \mathcal{L}(V)$, 并且 U 是 V 的子空间. 证明 U 在 T 下是不变的当且仅当 $P_U T P_U = T P_U$.

20. 设 $T \in \mathcal{L}(V)$, 并且 U 是 V 的子空间. 证明 U 和 $U^\perp$ 在 T 下都是不变的当且仅当 $P_U T = T P_U$.

21. 在 $\mathbf{R}^4$ 中, 设

$$U = \operatorname{span}((1,1,0,0),(1,1,1,2)).$$

求 $\boldsymbol{u} \in U$ 使得 $\|\boldsymbol{u} - (1,2,3,4)\|$ 最小.

22. 求 $\boldsymbol{p} \in \mathcal{P}_3(\mathbf{R})$, 使得 $\boldsymbol{p}(0) = 0$, $\boldsymbol{p}'(0) = 0$, 并且

$$\int_0^1 |2 + 3x - \boldsymbol{p}(x)|^2 \mathrm{d}x$$

最小.

23. 求 $\boldsymbol{p} \in \mathcal{P}_5(\mathbf{R})$ 使得

$$\int_{-\pi}^{\pi} |\sin x - \boldsymbol{p}(x)|^2 \mathrm{d}x$$

最小. (多项式 6.40 是本习题的一个极好的近似解, 但是这里要找的是包含 π 的幂的精确解. 可以使用能进行符号积分的电脑.)

24. 求多项式 $\boldsymbol{q} \in \mathcal{P}_2(\mathbf{R})$ 使得

$$\boldsymbol{p}\left(\frac{1}{2}\right) = \int_0^1 \boldsymbol{p}(x)\boldsymbol{q}(x)\mathrm{d}x, \quad \boldsymbol{p} \in \mathcal{P}_2(\mathbf{R}).$$

25. 求多项式 $\boldsymbol{q} \in \mathcal{P}_2(\mathbf{R})$ 使得

$$\int_0^1 \boldsymbol{p}(x)(\cos \pi x)\mathrm{d}x = \int_0^1 \boldsymbol{p}(x)\boldsymbol{q}(x)\mathrm{d}x, \quad \boldsymbol{p} \in \mathcal{P}_2(\mathbf{R}).$$