2.1.5 条件数与方程组的性态

考虑线性方程组

$$\begin{pmatrix} 2 & 6 \\ 2 & 6.00001 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 8.00001 \end{pmatrix}$$

它有准确解为： $x=(1,1)^T$ 。

如果方程组的系数矩阵以及右端项发生微小的变化，得

$$\begin{pmatrix} 2 & 6 \\ 2 & 5.99999 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 8.00002 \end{pmatrix}$$

它有准确解： $x=(10,-2)^{\mathrm{T}}$，可以看出，方程组的解变化非常大。

考虑线性方程组

$$\begin{pmatrix} 2 & 6 \\ 2 & 6.00001 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 8 \\ 8.00001 \end{pmatrix}$$

它有精确解为： $x=(1,1)^T$ 。

如果方程组的右端项发生微小的变化，$\delta \boldsymbol{b}=(0，0.00001)^T$ 得

$$\begin{pmatrix} 2 & 6 \\ 2 & 6.00001 \end{pmatrix}\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix}=\begin{pmatrix} 8 \\ 8.00002 \end{pmatrix}$$

其解为：$\tilde{x}=(-2,2)^{\mathrm{T}}$。这样，我们有

$$\frac{\|x-\tilde{x}\|_\infty}{\|x\|_\infty}=\frac{\left\|\begin{pmatrix} 3 \\ -1 \end{pmatrix}\right\|_\infty}{\left\|\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\|_\infty}=3,\quad \frac{\|\delta \boldsymbol{b}\|_\infty}{\|\boldsymbol{b}\|_\infty}=\frac{\left\|\begin{pmatrix} 0 \\ 0.00001 \end{pmatrix}\right\|_\infty}{\left\|\begin{pmatrix} 8 \\ 8.00001 \end{pmatrix}\right\|_\infty}=\frac{0.00001}{8.00001}\approx\frac{1}{800000},$$

即解的相对误差是右端的相对误差的**2400000**倍。

定义 如果线性方程组$Ax=b$中，A或 b的元素的微小变化，就会引起方程组解的巨大变化，则称方程组为“病态”方程组，矩阵A称为“病态”矩阵。否则称方程组为“良态”方程组，矩阵A称为“良态”矩阵。

我们需要一种能刻画矩阵和方程组“病态”标准的量。

求解$Ax=b$时，A 和b的误差对解x有何影响？

设A精确，b有误差 δb ，得到的解为 $x+\delta x$，即

$$\boxed{A(x+\delta x)=b+\delta b}$$

$$\cancel{Ax}+A\delta x=\cancel{b}+\delta b \Rightarrow \delta x = A^{-1}\delta b$$

$$\Rightarrow \qquad \|\delta x\| \le \|A^{-1}\|\cdot\|\delta b\|$$

又

$$\|b\|=\|Ax\|\le\|A\|\cdot\|x\| \Rightarrow \frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}$$

从而

相对误差放大因子

$$\boxed{\frac{\|\delta x\|}{\|x\|} \le \|A\|\cdot\|A^{-1}\|\cdot\frac{\|\delta b\|}{\|b\|}}$$

定义　设A为非奇异矩阵，$\|\cdot\|$为矩阵的算子范数，

则称 $\mathrm{cond}(A)=\|A\|\|A^{-1}\|$为矩阵$A$的条件数。

常用的条件数为：

$$\mathrm{cond}_\infty(A)=\|A\|_\infty\|A^{-1}\|_\infty$$

$$\mathrm{cond}_1(A)=\|A\|_1\|A^{-1}\|_1$$

$$\mathrm{cond}_2(A)=\|A\|_2\|A^{-1}\|_2=\sqrt{\frac{\lambda_{\max}(A^HA)}{\lambda_{\min}(A^HA)}}$$

分别称为矩阵A的∞–条件数、1–条件数和2–条件数。

注意，由 $A^H A = A^{-1} A A^H A = A^{-1}(AA^H)A$

得

$$\begin{aligned}\det(\lambda I - A^H A) &= \det(\lambda I - A^{-1}(AA^H)A) \\ &= \det(A^{-1}(\lambda I - (AA^H))A) \\ &= \det(A^{-1})\cdot\det(\lambda I - AA^H)\cdot\det(A) \\ &= \det(\lambda I - AA^H)\end{aligned}$$

则 $\lambda_{\max}(A^H A) = \lambda_{\max}(AA^H)$， $\|A\|_2^2 = \lambda_{\max}(A^H A)$

$$\begin{aligned}\|A^{-1}\|_2^2 &= \lambda_{\max}((A^{-1})^H A^{-1}) = \lambda_{\max}((A^{\mathrm{H}})^{-1} A^{-1}) \\ &= \lambda_{\max}((AA^{\mathrm{H}})^{-1}) = \lambda_{\max}((A^{\mathrm{H}}A)^{-1}) = \frac{1}{\lambda_{\min}(A^{\mathrm{H}}A)}\end{aligned}$$

故

$$\mathrm{cond}_2(A) = \|A\|_2 \|A^{-1}\|_2 = \sqrt{\frac{\lambda_{\max}(A^H A)}{\lambda_{\min}(A^H A)}}$$

矩阵的条件数具有如下的性质:

(1) $\mathrm{cond}(A) \geq 1$

$$\mathrm{cond}(A) = \left\|A^{-1}\right\|\left\|A\right\| \geq \left\|A^{-1}A\right\| = \left\|I\right\| = 1$$

(2) $\mathrm{cond}(A) = \mathrm{cond}(A^{-1})$

$$\mathrm{cond}(A^{-1}) = \left\|A^{-1}\right\| \cdot \left\|(A^{-1})^{-1}\right\| = \left\|A^{-1}\right\| \cdot \left\|A\right\| = \mathrm{cond}(A)$$

(3) $\mathrm{cond}(\alpha A) = \mathrm{cond}(A)$， $\alpha \neq 0$ ， $\alpha \in \mathbf{R}$

$$\mathrm{cond}(\alpha A) = \|\alpha A\| \cdot \|(\alpha A)^{-1}\| = |\alpha| \cdot \|A\| \cdot \frac{1}{|\alpha|} \cdot \|A^{-1}\|$$

$$= \|A\| \cdot \|A^{-1}\| = \mathrm{cond}(A)$$

(4) 如果U为酉（正交）矩阵，则

$$\mathrm{cond}_2(U) = 1$$

$$\mathrm{cond}_2(UA) = \mathrm{cond}_2(AU) = \mathrm{cond}_2(A)$$

由矩阵范数的等价性容易推出，$\mathbf{R}^{n\times n}$上任意两个范数的条件数$\mathbf{cond}_\alpha(A)$和$\mathbf{cond}_\beta(A)$都是等价性，即存在常数c_1和c_2，使得

$$c_1 \operatorname{cond}_\alpha(A) \le \operatorname{cond}_\beta(A) \le c_2 \operatorname{cond}_\alpha(A)$$

例如，有

$$\frac{1}{n}\operatorname{cond}_2(A) \le \operatorname{cond}_1(A) \le n \operatorname{cond}_2(A)$$

$$\frac{1}{n}\operatorname{cond}_\infty(A) \le \operatorname{cond}_2(A) \le n \operatorname{cond}_\infty(A)$$

$$\frac{1}{n^2}\operatorname{cond}_1(A) \le \operatorname{cond}_\infty(A) \le n^2 \operatorname{cond}_1(A)$$

这样，一个矩阵α范数下是病态的，则它在β范数下也是病态的。

$\mathrm{cond}(A)$ 越大，解的相对误差界可能越大，A对求解线性方程组来说就越可能呈现病态。

但$\mathrm{cond}(A)$多大A算病态，通常没有具体的定量标准； $\mathrm{cond}(A)$越小，解的相对误差界越小，反之，呈现病态。

n阶Hilbert矩阵

$$H_n=(h_{ij})_{n\times n}=\left(\frac{1}{i+j-1}\right)_{n\times n}=\begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{pmatrix} \quad i,j=1,2,\cdots,n$$

$$\mathrm{cond}(\boldsymbol{H}_4)=1.5514\times 10^4$$

$$\mathrm{cond}(\boldsymbol{H}_6)=1.4951\times 10^7$$

$$\mathrm{cond}(\boldsymbol{H}_8)=1.525\times 10^{10}$$

Hilbert矩阵常常出现在数据拟合和函数逼近的研究中。

与条件数有关的
一个数值例子、两个定理

一个数值例子：

在前面的例子中我们取 $\delta\boldsymbol{b}=(0,0.00001)^{T}$, $\delta A=O_{2\times2}$。

我们观察$\delta\boldsymbol{b}$对$\boldsymbol{x}$的影响，由

$$\boldsymbol{A}=\begin{pmatrix}2 & 6\\ 2 & 6.00001\end{pmatrix}$$ 易求， $$\boldsymbol{A}^{-1}=\begin{pmatrix}300000.5 & -300000\\ -100000 & 100000\end{pmatrix}$$

则$\boldsymbol{A}$的条件数为：

$$\operatorname{cond}_{\infty}(\boldsymbol{A})=\|\boldsymbol{A}\|_{\infty}\|\boldsymbol{A}^{-1}\|_{\infty}$$
$$=8.00001\times600000.5$$
$$\approx4800010\approx4.8\times10^{6}$$

则线性方程组的相对误差界为：

$$\frac{\|\delta \boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty} \le \mathrm{cond}_\infty(A)\frac{\|\delta \boldsymbol{b}\|_\infty}{\|\boldsymbol{b}\|_\infty} \approx 4.8\times 10^6 \times \frac{0.00001}{8}$$

$$\approx 4.8\times 10^6 \times 0.125\times 10^{-5}$$

$$\approx 6 \approx 600\%$$

可见，右端向量$\boldsymbol{b}$其分量百分之一的变化，可能引起解向量$\boldsymbol{x}$百分之六百的变化。

这说明矩阵A是严重病态矩阵，相应的线性方程组是病态方程组。

系数矩阵和右端项的扰动对解的影响

定理1 设 $Ax=b$，A 为非奇异矩阵，b为非零向量且 A 和 b 均有扰动。若A的扰动 δA 非常小，使得 $\|A^{-1}\|\|\delta A\|<1$，则

$$\frac{\|\delta x\|}{\|x\|}\le\frac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)\dfrac{\|\delta A\|}{\|A\|}}\left[\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right]$$

注：当 $\delta A=O$ 时，上述不等式为：$\dfrac{\|\delta x\|}{\|x\|}\le\mathrm{cond}(A)\dfrac{\|\delta b\|}{\|b\|}$

注：当 $\dfrac{\|\delta A\|}{\|A\|}$较小时，有 $\dfrac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)}\approx\mathrm{cond}(A)$，从而有

$$\frac{\|\delta x\|}{\|x\|}\lessapprox\mathrm{cond}(A)\left[\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right]$$

证 扰动后的方程组

$$(A+\delta A)(x+\delta x)=b+\delta b$$

注意到 $Ax=b$， $Ax+\delta Ax+\delta A\delta x+A\delta x=\delta b+b$

整理后，得

$$(A+\delta A)\,\delta x=\delta b-\delta Ax$$

由于A非奇异，故在δA充分小时，$A+\delta A$仍是非奇异的。事实上，由定理1.6知，只要$\|A^{-1}\|\,\|\delta A\|<1$， 就有$A+\delta A$可逆，而且

$$\left\|\left(I+A^{-1}\delta A\right)^{-1}\right\|\le\frac{1}{\left(1-\left\|A^{-1}\right\|\left\|\delta A\right\|\right)}$$

因此在此条件下，有 $(A+\delta A)=A\left(I+A^{-1}\delta A\right)$ 是非奇异的。而且

$$\begin{aligned}\delta x&=(A+\delta A)^{-1}\left(\delta b-\delta Ax\right)\\&=(I+A^{-1}\delta A)^{-1}A^{-1}\left(\delta b-\delta Ax\right)\end{aligned}$$

将上式两端取范数，

进一步,

$$\frac{1}{\|x\|}\|\delta x\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|\delta A\|}\left(\|\delta b\|+\|\delta A\|\cdot\|x\|\right)\frac{1}{\|x\|}$$

再利用 $\frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}$ 有

$$\frac{\|\delta x\|}{\|x\|} \le \frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|\delta A\|}\left(\|\delta b\|\frac{\|A\|}{\|b\|}+\frac{\|\delta A\|\cdot\|x\|}{\|x\|}\frac{\|A\|}{\|A\|}\right)$$

$$=\frac{\|A^{-1}\|\cdot\|A\|}{1-\|A^{-1}\|\cdot\|\delta A\|}\left(\frac{\|\delta b\|}{\|b\|}+\frac{\|\delta A\|}{\|A\|}\right)$$

$$=\frac{\|A^{-1}\|\cdot\|A\|}{1-\|A^{-1}\|\cdot\|A\|\cdot\frac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta b\|}{\|b\|}+\frac{\|\delta A\|}{\|A\|}\right)=\frac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)\frac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta b\|}{\|b\|}+\frac{\|\delta A\|}{\|A\|}\right)$$

条件数的几何意义

定理 设$A\in\mathbf{R}^{n\times n}$非奇异，则

$$\min\left\{\frac{\|\delta A\|_2}{\|A\|_2}:\quad A+\delta A\ \text{奇异}\right\}=\frac{1}{\|A^{-1}\|_2\cdot\|A\|_2}=\frac{1}{\mathbf{cond}_2(A)}$$

即在谱范数下，一个矩阵的条件数的倒数正好等于该矩阵与全体奇异矩阵所成集合的相对距离。

定理表明，当$A\in\mathbf{R}^{n\times n}$十分病态时，就说明$A$已与一个奇异矩阵十分接近。

近似解的余量与它的相对误差间的关系

定理2 设$Ax=b$，A为非奇异矩阵，b为非零向量，则方程组近似解$\tilde{x}$的事后估计式为

$$\frac{1}{\mathrm{cond}(A)}\frac{\|b-A\tilde{x}\|}{\|b\|}\le\frac{\|\tilde{x}-x\|}{\|x\|}\le\mathrm{cond}(A)\frac{\|b-A\tilde{x}\|}{\|b\|}$$

其中称$\|b-A\tilde{x}\|$为近似解$\tilde{x}$的余量，简称余量。

证　由$Ax=b$，得 $b-A\tilde{x}=Ax-A\tilde{x}=A(x-\tilde{x})$　故

$(x-\tilde{x})=A^{-1}(b-A\tilde{x})$，利用 $\|x-\tilde{x}\|\le\|A^{-1}\|\cdot\|b-A\tilde{x}\|$，$\dfrac{1}{\|x\|}\le\dfrac{\|A\|}{\|b\|}$

$$\frac{\|\tilde{x}-x\|}{\|x\|}\le\frac{\|A\|}{\|b\|}\|A^{-1}\|\|b-A\tilde{x}\|=\mathrm{cond}(A)\frac{\|b-A\tilde{x}\|}{\|b\|}$$

由$x=A^{-1}b$，$\|x\|\le\|A^{-1}\|\cdot\|b\|$，再利用 $\|x-\tilde{x}\|\|A\|\ge\|b-A\tilde{x}\|$，得

$$\frac{\|b-A\tilde{x}\|}{\|A^{-1}\|\|b\|}\le\frac{\|b-A\tilde{x}\|}{\|x\|}\le\|A\|\frac{\|\tilde{x}-x\|}{\|x\|}$$，即 $$\frac{1}{\mathrm{cond}(A)}\frac{\|b-A\tilde{x}\|}{\|b\|}\le\frac{\|\tilde{x}-x\|}{\|x\|}$$

该定理表明，只要当**cond**$(A)\approx1$时，近似解余量的相对误差是解的相对误差的一个好的度量。当**cond**(A)很大，若方程组呈现病态时，虽然近似解余量的相对误差已经很小，但解的相对误差仍然很大。

返回本节

2.1.5 矩阵的QR分解

如何利用直接法求解一些病态方程组？

Gauss消去过程实际上是用一系列具有特定结构的单位下三角矩阵将$\boldsymbol{A}$逐步上三角化的过程。由矩阵的条件数定义可以看出，正交矩阵是性态最好的矩阵，如果我们能用正交矩阵代替Gauss消去过程中的单位下三角矩阵，即

$$
A=\begin{pmatrix}\times & \times & \times\\ \times & \times & \times\\ \times & \times & \times\end{pmatrix}\xrightarrow{\text{用正交变换}\boldsymbol{Q}_2}\begin{pmatrix}\times & \times & \times\\ 0 & \times & \times\\ 0 & 0 & \times\end{pmatrix}=\boldsymbol{U}
$$

则$Q_1Q_2A=U$，计算知$\mathrm{cond}_2(A)=\mathrm{cond}_2(U)$。因此变换后所得的矩阵$U$的条件数不变， 故该计算过程具有数值稳定性。

为实现矩阵一般的QR分解，我们引入矩阵

Householder变换矩阵

定义2.4 设$\boldsymbol{\omega} \in \mathbf{R}^n, \boldsymbol{\omega} \neq 0$，称初等矩阵

$$\boldsymbol{H}(\boldsymbol{\omega}) = \boldsymbol{I} - \frac{2}{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}}\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}$$

为**Householder矩阵**（简称$\boldsymbol{H}$阵），或称其为**Householder变换**。

Householder矩阵是一种特殊的初等变换矩阵。

一般的初等矩阵表示为：

$$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\alpha)=\boldsymbol{I}-\alpha\,\boldsymbol{u}\boldsymbol{v}^{H},\ \ \boldsymbol{u},\boldsymbol{v}\in\mathbf{C}^{\mathbf{n}},\ \alpha\in\mathbf{C}$$

此时， $\boldsymbol{u}\boldsymbol{v}^{H}$ 是一个秩1矩阵，即

$$1\le\ \mathbf{rank}(\boldsymbol{u}\boldsymbol{v}^{H})\le\mathbf{rank}(\boldsymbol{u})=1$$

其特征值为： $\boldsymbol{v}^{H}u,\ \underbrace{0,\cdots,0}_{n-1}$

事实上， $\left(\boldsymbol{u}\boldsymbol{v}^{H}\right)\boldsymbol{u}=\boldsymbol{u}\left(\boldsymbol{v}^{H}\boldsymbol{u}\right)=\left(\boldsymbol{v}^{H}\boldsymbol{u}\right)\boldsymbol{u}$

即 $\lambda=\boldsymbol{v}^{H}\boldsymbol{u}\neq0$ 是矩阵 $\boldsymbol{u}\boldsymbol{v}^{H}$ 的唯一非零特征值。

特别取 设 $\boldsymbol{\omega}=\left(\boldsymbol{e}_i-\boldsymbol{e}_j\right)$ ，则有

$$\boldsymbol{H}(\boldsymbol{\omega})=\boldsymbol{I}-\frac{2}{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}}\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}$$

$$=\boldsymbol{I}-\left(\boldsymbol{e}_i-\boldsymbol{e}_j\right)\left(\boldsymbol{e}_i-\boldsymbol{e}_j\right)^T$$

其中

$$\boldsymbol{e}_i=(0,0,\cdots,0,\underset{\text{第 } i \text{ 个分量}}{1},0,\cdots,0)^T$$

$$\boldsymbol{e}_j=(0,0,\cdots,0,\underset{\text{第 } j \text{ 个分量}}{1},0,\cdots,0)^T$$

则为初等置换矩阵$\boldsymbol{P}_{ij}$。

初等置换矩阵是特殊的$\boldsymbol{H}$矩阵。

显然Householder矩阵矩阵具有如下性质：

(1) $H(\omega)^{\mathrm{T}}=H(\omega)$，即 H 阵为对称阵；

$$H(\omega)^T=\left(I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}\right)^T=I-\frac{2}{\omega^{\mathrm{T}}\omega}\left(\omega\omega^{\mathrm{T}}\right)^T=I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}$$

(2) $H(\omega)^{\mathrm{T}}H(\omega)=I_n$，即$H$ 阵为正交阵；

$$H(\omega)^T H(\omega)=\left(I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}\right)^2$$

$$=I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}+\left(\frac{2}{\omega^{\mathrm{T}}\omega}\right)^2\left(\omega\omega^{\mathrm{T}}\right)\left(\omega\omega^{\mathrm{T}}\right)$$

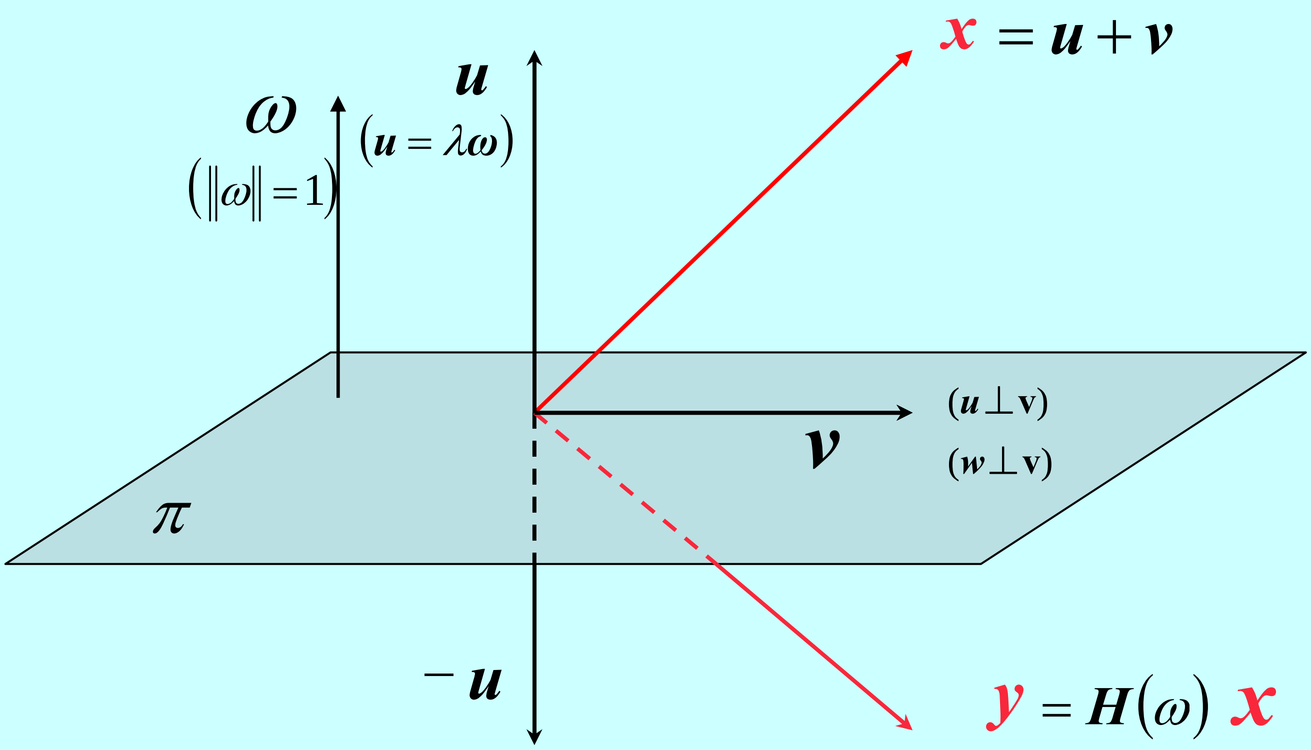
$$=I-\frac{4}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}+\frac{4}{\left(\omega^{\mathrm{T}}\omega\right)^2}\left(\omega^{\mathrm{T}}\omega\right)\left(\omega\omega^{\mathrm{T}}\right)=I$$

(3) 如果$H(\omega)x = y$，则$\|y\|_2 = \|x\|_2$； 事实上，

$$\begin{aligned}\|y\|_2^2 = y^T y &= (H(\omega)x)^T (H(\omega)x) \\ &= x^T\left(H(\omega)^T H(\omega)\right)x \\ &= x^T x = \|x\|_2^2\end{aligned}$$

(亦称Householder矩阵为镜面反射变换)；

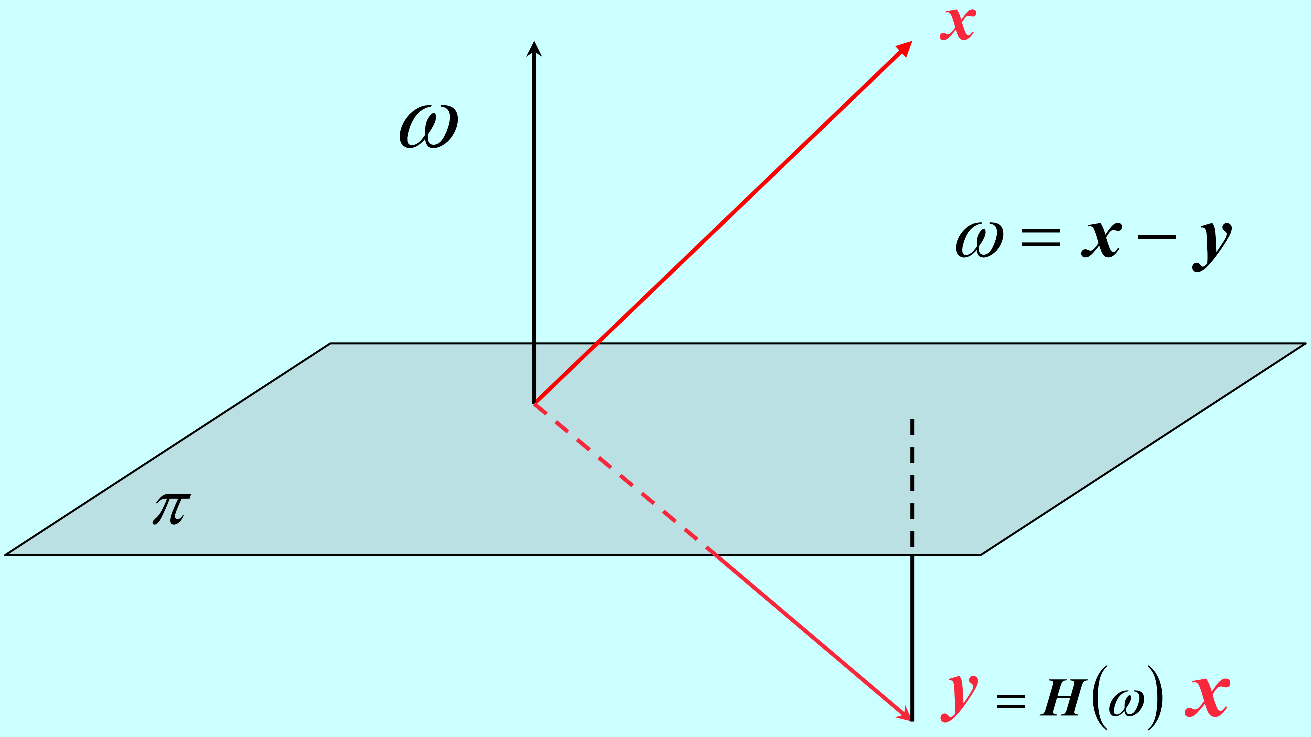
性质3亦称镜面反射变换，其几何意义如下：



$$H(\omega)x = \left(I - 2\omega\omega^T\right)x = x - 2\omega\omega^T(u+v) = x - 2\omega\left(\omega^T u + \omega^T v\right)$$

$$= u + v - 2\lambda\omega\left(\omega^T\omega\right) = u + v - 2u \quad = -u + v$$

将向量x变成通长已知向量y的Householder矩阵中ω的选取方法:



例，求将向量$\boldsymbol{x}=(3, 4)^T$映射成为$\boldsymbol{y}=(4, 3)^T$的正交变换阵$\boldsymbol{H}$。

解：取$\omega=\boldsymbol{x}-\boldsymbol{y}=(-1, 1)^T$，则 $\boldsymbol{\omega}^T\boldsymbol{\omega}=2$

$$\boldsymbol{H}(\boldsymbol{\omega})=\boldsymbol{I}-\frac{2}{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}}\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}=\begin{pmatrix}1&0\\0&1\end{pmatrix}-\begin{pmatrix}-1\\1\end{pmatrix}\begin{pmatrix}-1&1\end{pmatrix}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$$

例，求将向量$\boldsymbol{x}=(3, 4)^T$映射成为$\boldsymbol{y}=(0, 5)^T$的正交变换阵$\boldsymbol{H}$。

解：取$\omega=\boldsymbol{x}-\boldsymbol{y}=(3, -1)^T$，则 $\boldsymbol{\omega}^T\boldsymbol{\omega}=10$

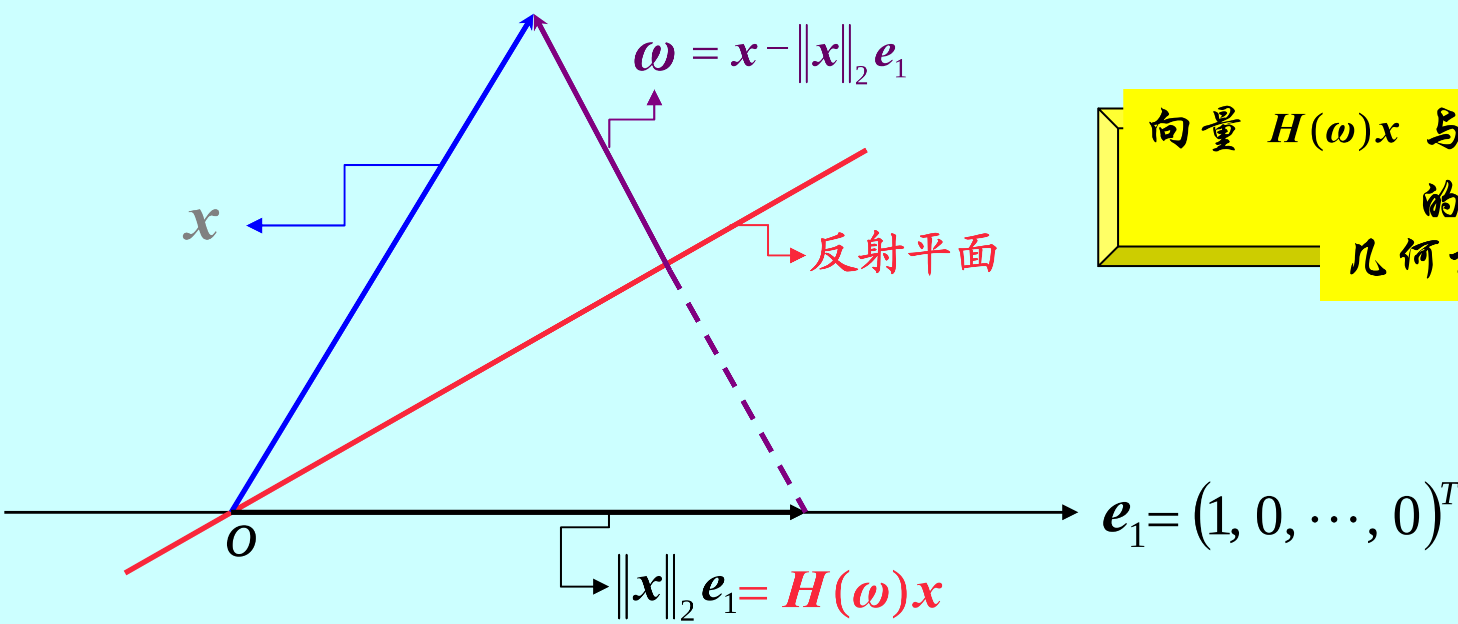
$$\boldsymbol{H}(\boldsymbol{\omega})=\boldsymbol{I}-\frac{2}{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}}\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}=\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{1}{5}\begin{pmatrix}3\\-1\end{pmatrix}\begin{pmatrix}3&-1\end{pmatrix}$$

$$=\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{1}{5}\begin{pmatrix}9&-3\\-3&1\end{pmatrix}=\begin{pmatrix}-\frac{4}{5}&\frac{3}{5}\\\frac{3}{5}&\frac{4}{5}\end{pmatrix}$$

(4) 设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{R}^n$ 且 $\boldsymbol{x}\neq\boldsymbol{0}$, 取 $\boldsymbol{\omega}=\boldsymbol{x}-\|\boldsymbol{x}\|_2\boldsymbol{e}_1$

则

$$\boldsymbol{H}(\boldsymbol{\omega})\boldsymbol{x}=\boldsymbol{H}\left(\boldsymbol{x}-\|\boldsymbol{x}\|_2\boldsymbol{e}_1\right)\boldsymbol{x}=\left(\|\boldsymbol{x}\|_2,0,\cdots,0\right)^T=\|\boldsymbol{x}\|_2\boldsymbol{e}_1$$



向量 $H(\omega)x$ 与向量 e_1 共线的
几何证明

注释: $Ax = b \Leftrightarrow (QR)x = b \Leftrightarrow Rx = Q^T b$

$$\begin{cases} 1.\ \text{计算}QR\text{因子分解}A=QR \\ 2.\ \text{计算}y=Q^T b \\ 3.\ \text{对}x\text{解}Rx=y \end{cases}$$

这种直接法的数值稳定性要比LU分解好，但是QR分解的计算量远远大于LU分解，因此，QR分解只适用于求解病态线性方程组。

下面我们利用一系列H阵将A分解成QR形式。

例 利用Householder变换求A的QR分解，其中

$$A=\begin{pmatrix}0&3&1\\0&4&-2\\2&1&2\end{pmatrix}.$$

解： 将A按列分块为 $A=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$,其中 $\boldsymbol{a}_1=(0,0,2)^{\mathrm{T}}$,

$\|\boldsymbol{a}_1\|_2=2$, 取 $\omega_1=\boldsymbol{a}_1-\|\boldsymbol{a}_1\|_2\boldsymbol{e}_1=\begin{pmatrix}-2\\0\\2\end{pmatrix}$，则

$$\omega_1^T\omega_1=\begin{pmatrix}-2&0&2\end{pmatrix}\begin{pmatrix}-2\\0\\2\end{pmatrix}=8,$$ 令

$$Q_1 = H(\omega_1) = I - \frac{2}{\omega_1^{\mathrm{T}}\omega_1}\omega_1\omega_1^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{4}\begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}\begin{pmatrix} -2 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$Q_1 A = H(\omega_1)A = (H(\omega_1)a_1, H(\omega_1)a_2, H(\omega_1)a_3)$$

$$= \begin{pmatrix} 2 & 1 & 2 \\ 0 & 4 & -2 \\ 0 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 2 & \boldsymbol{b}^{\mathrm{T}} \\ \boldsymbol{0} & A_2 \end{pmatrix}$$

其中

$$\boldsymbol{b}^{\mathrm{T}}=(1,2),\quad \boldsymbol{A}_2=\begin{pmatrix}4 & -2\\ 3 & 1\end{pmatrix}=(\tilde{\boldsymbol{a}}_1,\tilde{\boldsymbol{a}}_2)$$

$\tilde{\boldsymbol{a}}_1=(4,3)^{\mathrm{T}}$, $\tilde{\boldsymbol{a}}_2=(-2,1)^{\mathrm{T}}$, $\|\tilde{\boldsymbol{a}}_1\|_2=5$, 取

$$\boldsymbol{\omega}_2=\tilde{\boldsymbol{a}}_1-\|\tilde{\boldsymbol{a}}_1\|_2\boldsymbol{e}_1=\begin{pmatrix}4\\3\end{pmatrix}-\begin{pmatrix}5\\0\end{pmatrix}=\begin{pmatrix}-1\\3\end{pmatrix}$$

$$\boldsymbol{\omega}_2^T\boldsymbol{\omega}_2=\|\boldsymbol{\omega}_2\|_2^2=10$$

$$H(\omega_2)=\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{1}{5}\begin{pmatrix}-1\\3\end{pmatrix}\begin{pmatrix}-1&3\end{pmatrix}$$

$$=\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{1}{5}\begin{pmatrix}1&-3\\-3&9\end{pmatrix}=\frac{1}{5}\begin{pmatrix}4&3\\3&-4\end{pmatrix}$$

$$H(\omega_2)A_2=\left(H(\omega_2)\tilde{a}_1,H(\omega_2)\tilde{a}_2\right)$$

$$=\left[\frac{1}{5}\begin{pmatrix}4&3\\3&-4\end{pmatrix}\begin{pmatrix}4\\3\end{pmatrix},\frac{1}{5}\begin{pmatrix}4&3\\3&-4\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix}\right]=\begin{pmatrix}5&-1\\0&-2\end{pmatrix}。$$

令

$$\boldsymbol{Q}_2=\begin{pmatrix}1 & \boldsymbol{0}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{H}(\omega_2)\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & \dfrac{4}{5} & \dfrac{3}{5}\\ 0 & \dfrac{3}{5} & -\dfrac{4}{5}\end{pmatrix}$$

则

$$(\boldsymbol{Q}_2\boldsymbol{Q}_1)\boldsymbol{A}=\begin{pmatrix}1 & \boldsymbol{0}^{T}\\ \boldsymbol{0} & \boldsymbol{H}(\omega_2)\end{pmatrix}\begin{pmatrix}2 & \boldsymbol{b}^{T}\\ \boldsymbol{0} & \boldsymbol{A}_2\end{pmatrix}=\begin{pmatrix}2 & \boldsymbol{b}^{T}\\ \boldsymbol{0} & \end{pmatrix}$$

$$=\begin{pmatrix}2 & 1 & 2\\ 0 & 5 & -1\\ 0 & 0 & -2\end{pmatrix}=\boldsymbol{R}$$

令
$$\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{Q}_2\boldsymbol{Q}_1=\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 0 & \frac{3}{5} & -\frac{4}{5}\end{pmatrix}\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}=\begin{pmatrix}0 & 0 & 1\\ \frac{3}{5} & \frac{4}{5} & 0\\ -\frac{4}{5} & \frac{3}{5} & 0\end{pmatrix}$$

从而得到$\boldsymbol{Q}$和$\boldsymbol{A}$的$\boldsymbol{QR}$分解如下：

$$\boldsymbol{Q}=\begin{pmatrix}0 & \frac{3}{5} & -\frac{4}{5}\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 1 & 0 & 0\end{pmatrix}\qquad \boldsymbol{A}=\boldsymbol{QR}=\begin{pmatrix}0 & \frac{3}{5} & -\frac{4}{5}\\ 0 & \frac{4}{5} & \frac{3}{5}\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}2 & 1 & 2\\ 0 & 5 & -1\\ 0 & 0 & -2\end{pmatrix}。$$

例 利用**Householder**变换求A的QR分解，其中

$$A=\begin{pmatrix}0&4\\0&0\\5&2\end{pmatrix}.$$

解法之一：

解 首先将A按列分块为$A=(a_1,a_2)$，其中$a_1=(0,0,5)^{\mathrm{T}}$，则$\|a_1\|_2=5$

取

$$\omega_1=a_1-\|a_1\|_2e_1=\begin{pmatrix}-5\\0\\5\end{pmatrix}，则\ \omega_1^T\omega_1=(-5\quad 0\quad 5)\begin{pmatrix}-5\\0\\5\end{pmatrix}=5,$$

$$Q_1=H(\omega_1)=I-\frac{2}{\omega_1^{\mathrm{T}}\omega_1}\omega_1\omega_1^{\mathrm{T}}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\frac{1}{25}\begin{pmatrix}-5\\0\\5\end{pmatrix}(-5\quad 0\quad 5)=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\frac{1}{25}\begin{pmatrix}25&0&-25\\0&0&0\\-25&0&25\end{pmatrix}$$

$$Q_1 A = H(\omega_1)A = (H(\omega_1)a_1, H(\omega_1)a_2) = \begin{pmatrix} 5 & 2 \\ 0 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ \boldsymbol{0} & A_2 \end{pmatrix},$$ 其中

$$A_2 = \begin{pmatrix} 0 \\ 4 \end{pmatrix} = (\tilde{a}_1),$$ 则 $\|\tilde{a}_1\|_2 = 4$，取 $\omega_2 = \tilde{a}_1 - \|\tilde{a}_1\|_2 e_1 = \begin{pmatrix} 0 \\ 4 \end{pmatrix} - \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} -4 \\ 4 \end{pmatrix}$

$$\omega_2^T \omega_2 = \|\omega_{21}\|_2^2 = 32, \quad H(\omega_2) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} - \frac{1}{16}\begin{pmatrix} 16 & -16 \\ -16 & 16 \end{pmatrix}$$

那么 $\begin{pmatrix} -4 & 4 \end{pmatrix}$ 有

$$H(\omega_2)A_2 = H(\omega_2)\tilde{a}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix},$$ 令

$$Q_2 = \begin{pmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \boldsymbol{0} & H(\omega_2) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$ ，则

$$(Q_2Q_1)A=\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & H(\omega_2)\end{pmatrix}\begin{pmatrix}5 & 0\\ \boldsymbol{0} & A_2\end{pmatrix}=\begin{pmatrix}5 & 0\\ 0 & \end{pmatrix}=\begin{pmatrix}5 & 2\\ 0 & 4\\ 0 & 0\end{pmatrix}=\begin{pmatrix}\bar{\boldsymbol{R}}\\ \boldsymbol{0}\end{pmatrix}$$

$$\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{Q}_2\boldsymbol{Q}_1=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$$ ，则

$$\boldsymbol{Q}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$$ ，从而得 $A=QR$ ，即

$$\begin{pmatrix}0&4\\0&0\\5&2\end{pmatrix}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\begin{pmatrix}5&2\\0&4\\0&0\end{pmatrix}$$