第 3 章

最小二乘问题

3.1 引言

最小二乘问题（Least Squares）包括线性最小二乘问题、总体最小二乘问题、等式约束最小二乘问题、刚性加权最小二乘问题等，它在最优化问题、材料与结构力学、信号与图像处理等方面有广泛应用。

本章我们考虑线性最小二乘问题

$$\min_{x\in R^n}\|\boldsymbol{r}(\boldsymbol{x})\|_2=\min_{x\in R^n}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2 \tag{3.1.1}$$

的求解，其中，$\boldsymbol{A}\in R^{m\times n}$；$\boldsymbol{b}\in R^m$。通常我们称式（3.1.1）的解为最小二乘解，其中，$\boldsymbol{r}(\boldsymbol{x})$ 常常被称为残量。

- 当 $m=n$ 且 $\boldsymbol{A}$ 非奇异时，式（3.1.1）就是一个线性方程组，解为 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$；
- 当 $m>n$ 时，约束个数大于未知量个数，此时我们称式（3.1.1）为超定的（overdetermined）；
- 当 $m<n$ 时，未知量个数大于约束个数，此时我们称式（3.1.1）为欠定的（underdetermined）。

下面介绍一些符号和概念。

设矩阵 $\boldsymbol{A}\in R^{m\times n}$，则 $\boldsymbol{A}$ 的值域定义为

$$\mathrm{Ran}(\boldsymbol{A})=\left\{\boldsymbol{y}\in R^m : \boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}\in R^n\right\}$$

$\boldsymbol{A}$ 的零空间定义为

$$\mathrm{Ker}(\boldsymbol{A})=\left\{\boldsymbol{x}\in R^n : \boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\right\}$$

一个子空间的正交补的定义为

$$S^{\perp}=\left\{\boldsymbol{y}\in R^n : \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}, \forall \boldsymbol{x}\in S\right\}$$

Moore-Penrose 广义逆的定义为：若矩阵 $\boldsymbol{X}\in R^{m\times n}$ 满足

$$\boldsymbol{AXA}=\boldsymbol{A}，\quad \boldsymbol{XAX}=\boldsymbol{X}，\quad (\boldsymbol{AX})^{*}=\boldsymbol{AX}，\quad (\boldsymbol{XA})^{*}=\boldsymbol{XA}$$

则称 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的 Moore-Penrose 广义逆，通常记作 $\boldsymbol{A}^{\dagger}$ 。

定理 3.1.1　设 $\boldsymbol{A},\boldsymbol{E}\in\mathbb{C}^{m\times n}$（$m\geqslant n$）则 $\boldsymbol{x}_*\in R^n$ 是式（3.1.1）线性最小二乘问题的解当且仅当残量 $\boldsymbol{r}=\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_*$ 与 $\mathrm{Ran}(\boldsymbol{A})$（值域）正交，即 $\boldsymbol{x}_*$ 是下面正规方程的解

$$\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x})=\boldsymbol{0} \text{ 或 } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \tag{3.1.2}$$

证明　充分性：设 $\boldsymbol{x}_*$ 是正规方程（3.1.2）的解。对任意向量 $\boldsymbol{y}\in\mathbb{R}^n$，由 $(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_*)\perp\mathrm{Ran}(\boldsymbol{A})$ 可知

$$\begin{aligned}\|\boldsymbol{A}\boldsymbol{y}-\boldsymbol{b}\|_2^2&=\left\|(\boldsymbol{A}\boldsymbol{x}_*-\boldsymbol{b})+\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{x}_*)\right\|_2^2\\&=\|\boldsymbol{A}\boldsymbol{x}_*-\boldsymbol{b}\|_2^2+\left\|\boldsymbol{A}(\boldsymbol{y}-\boldsymbol{x}_*)\right\|_2^2\end{aligned}$$

因此，$\boldsymbol{x}_*$ 是式（3.1.1）线性最小二乘问题的解。

必要性：设 $\boldsymbol{x}_*$ 是式（3.1.1）线性最小二乘问题的解。用反证法，假定 $\boldsymbol{z}\triangleq\boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_*\right)\neq\boldsymbol{0}$ 。取 $\boldsymbol{y}=\boldsymbol{x}_*+\alpha\boldsymbol{z}$，其中，$\alpha>0$ 。则有

$$\|\boldsymbol{A}\boldsymbol{y}-\boldsymbol{b}\|_2^2=\|\boldsymbol{A}\boldsymbol{x}_*-\boldsymbol{b}+\alpha\boldsymbol{A}\boldsymbol{z}\|_2^2=\|\boldsymbol{A}\boldsymbol{x}_*-\boldsymbol{b}\|_2^2-2\alpha\|\boldsymbol{z}\|_2^2+\alpha^2\|\boldsymbol{A}\boldsymbol{z}\|_2^2$$

由于 $\|\boldsymbol{z}\|_2>0$，所以当 α 充分小时，有 $2\|\boldsymbol{z}\|_2^2>\alpha\|\boldsymbol{A}\boldsymbol{z}\|_2^2$，即上式右端小于 $\|\boldsymbol{A}\boldsymbol{x}_*-\boldsymbol{b}\|_2^2$。这与 $\boldsymbol{x}_*$ 是最小二乘解相矛盾。所以 $\boldsymbol{z}=\boldsymbol{0}$，即 $\boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_*\right)=\boldsymbol{0}$ 。

由定理 3.1.1　可知，求式（3.1.1）线性最小二乘问题的解等价于求正规方程（3.1.2）的解。由于

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\in\mathrm{Ran}\left(\boldsymbol{A}^{\mathrm{T}}\right)=\mathrm{Ran}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$$

所以正规方程 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 是相容（consistent）的，即最小二乘解总是存在的。当 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 非奇异时，这个解是唯一的。

定理 3.1.2　设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$（$m>n$），则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 对称正定当且仅当 $\boldsymbol{A}$ 是列满秩的，即 $\mathrm{rank}(\boldsymbol{A})=n$。

此时，式（3.1.1）线性最小二乘问题的解是唯一的，其表达式为

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

若定义 $\boldsymbol{A}^{\dagger}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}$，则式（3.1.1）线性最小二乘问题的解可表示为

$$\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{b}$$

下面介绍最小二乘解的几何含义。

根据定理 3.1.1，我们可以把 $\boldsymbol{b}$ 写成

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_* + \boldsymbol{r}\text{，其中，}\boldsymbol{r} \perp \mathrm{Ran}(\boldsymbol{A}) \tag{3.1.3}$$

所以 $\boldsymbol{A}\boldsymbol{x}_*$ 就是 $\boldsymbol{b}$ 在 Ran($\boldsymbol{A}$) 上的正交投影（见图 3.1.1），最小二乘解可能并不唯一，但分解式（3.1.3）总是唯一的。

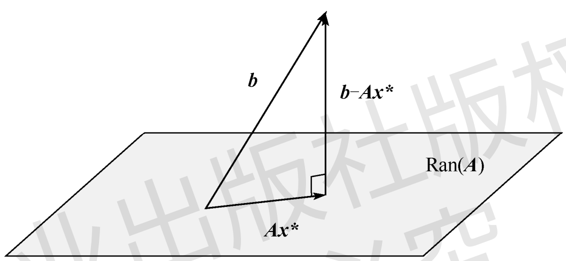

图 3.1.1　最小二乘解的几何描述

线性最小二乘问题的一个重要应用是低次多项式数据拟合，即已知平面上 n 个点 $\{(t_i, f_i)\}_{i=1}^n$ 寻找一个低次多项式来拟合这些数据。不妨设拟合多项式为

$$p(x) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m$$

通常 $m \ll n$。将 n 个点 $\{(t_i, f_i)\}_{i=1}^n$ 代入上式可得

$$\begin{bmatrix} 1 & t_1 & t_1^2 & t_1^m \\ 1 & t_2 & t_2^2 & t_2^m \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & t_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \text{或 } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{f} \tag{3.1.4}$$

其中，$\boldsymbol{A} \in \mathbb{R}^{n\times(m+1)}, \boldsymbol{x} = [a_0, a_1, \cdots, a_m]^{\mathrm{T}}$。当 $m = n$ 时，求解方程组（3.1.4）可利用第 2 章的方法去讨论。当 $m \ll n$ 时，该方程组是超定的，解通常不存在。因此，我们需要寻找一个近似解，使得残量 $\|\boldsymbol{f} - \boldsymbol{A}\boldsymbol{x}\|_2$ 最小，即求解最小二乘问题

$$\min_{\boldsymbol{x}\in\mathbb{R}^{m+1}} \|\boldsymbol{f} - \boldsymbol{A}\boldsymbol{x}\|_2^2$$

由于篇幅有限，本章主要讨论超定线性最小二乘问题的性质与求解。在此之前，我们先给出一些理论工具。

3.2　初等变换矩阵

初等变换矩阵计算的一个基本思想就是把较复杂的问题简易化。常用的初等变换矩阵有 3 类：Gauss 变换、Householder 变换和 Givens 变换。

3.2.1　初等矩阵

考虑初等矩阵

$$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})=\boldsymbol{I}-\tau\boldsymbol{u}\boldsymbol{v}^*$$

其中，$\boldsymbol{u},\boldsymbol{v}\in C^n$ 是非零向量，$\boldsymbol{\tau}$ 是一个非零复数。事实上，$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})$ 是单位矩阵的一个秩 1 扰动。

定理 3.2.1　设 $\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})$ 是一个初等矩阵，我们有

（1） $\det\left[\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})\right]=1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}$；

（2）若 $1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}\neq 0$，则 $\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})$ 非奇异，且 $\left[\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})\right]^{-1}=\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\gamma})$，其中，$\boldsymbol{\gamma}=\dfrac{\boldsymbol{\tau}}{\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}-1}$。

证明　（1）不难验证

$$\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{v}^* & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{I}\text{-}\boldsymbol{\tau}\boldsymbol{u}\boldsymbol{v}^* & -\boldsymbol{\tau}\boldsymbol{u}\\ \boldsymbol{0} & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ -\boldsymbol{v}^* & 1\end{bmatrix}=\begin{bmatrix}\boldsymbol{I} & -\boldsymbol{\tau}\boldsymbol{u}\\ \boldsymbol{v}^* & 1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}\end{bmatrix}$$

由行列式的乘法可知

$$\det\left(\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{v}^* & 1\end{bmatrix}\right)\det\left(\begin{bmatrix}\boldsymbol{I}-\boldsymbol{\tau}\boldsymbol{u}\boldsymbol{v}^* & -\boldsymbol{\tau}\boldsymbol{u}\\ \boldsymbol{0} & 1\end{bmatrix}\right)\det\left(\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0}\\ -\boldsymbol{v}^* & 1\end{bmatrix}\right)=\det\left(\begin{bmatrix}\boldsymbol{I} & -\boldsymbol{\tau}\boldsymbol{u}\\ \boldsymbol{v}^* & 1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}\end{bmatrix}\right)$$

所以有 $\det\left[\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})\right]=\det\left(1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}\right)=1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}$。

（2）若 $1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}\neq 0$，则 $\det\left[\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})\right]\neq 0$，即 $\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})$ 非奇异，所以

$$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau})\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{\gamma})=\boldsymbol{I}-\boldsymbol{\tau}\boldsymbol{u}\boldsymbol{v}^*-\boldsymbol{\gamma}(1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u})\boldsymbol{u}\boldsymbol{v}^*=\boldsymbol{I}-\boldsymbol{\tau}\boldsymbol{u}\boldsymbol{v}^*-\frac{\boldsymbol{\tau}}{\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u}-1}(1-\boldsymbol{\tau}\boldsymbol{v}^*\boldsymbol{u})\boldsymbol{u}\boldsymbol{v}^*=\boldsymbol{I}$$

定理 3.2.2　设 $\boldsymbol{A}\in C^{n\times n}$，则 $\boldsymbol{A}$ 非奇异当且仅当 $\boldsymbol{A}$ 可以分解成若干个初等矩阵的乘积。

3.2.2 Gauss 变换

设 $\boldsymbol{l}_j=[0,\cdots,0,l_{j+1,j},\cdots,l_{n,j}]^{\mathrm{T}}$，$j=1,2,\cdots,n$，则 Gauss 变换定义为

$$\boldsymbol{L}(\boldsymbol{l}_j)=\boldsymbol{E}(\boldsymbol{l}_j,\boldsymbol{e}_j,-1)=\boldsymbol{I}+\boldsymbol{l}_j\boldsymbol{e}_j^{\mathrm{T}}=\begin{bmatrix}1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & l_{j+1,j} & 1 & \\ & & \vdots & & \ddots \\ & & l_{n,j} & & & 1\end{bmatrix}$$

向量 $\boldsymbol{l}_j$ 称为 Gauss 向量。由定理 3.2.1 可知

$$\det\left[\boldsymbol{L}(\boldsymbol{l}_j)\right]=1,\ \left[\boldsymbol{L}(\boldsymbol{l}_j)\right]-1=\boldsymbol{E}(\boldsymbol{l}_j,\boldsymbol{e}_j,1)=\boldsymbol{E}(-\boldsymbol{l}_j,\boldsymbol{e}_j,-1)=\boldsymbol{L}(-\boldsymbol{l}_j)$$

3.2.3 Householder 变换

定义 3.2.1 我们称矩阵

$$\boldsymbol{H}=\boldsymbol{I}-\frac{2}{\boldsymbol{v}^*\boldsymbol{v}}\boldsymbol{v}\boldsymbol{v}^*=\boldsymbol{I}-\frac{2}{\|\boldsymbol{v}\|_2^2}\boldsymbol{v}\boldsymbol{v}^*,\ \boldsymbol{0}\neq\boldsymbol{v}\in\mathbb{C}^n \tag{3.2.1}$$

为 Householder 矩阵（或 Householder 变换，或 Householder 反射），向量 $\boldsymbol{v}$ 称为 Householder 向量。我们通常将式（3.2.1）记为 $\boldsymbol{H}(\boldsymbol{v})$。

从几何上看，一个 Householder 变换就是一个关于超平面 $\mathrm{span}\{\boldsymbol{v}\}^{\perp}$ 的反射。对任意一个向量 $\boldsymbol{x}\in C^n$，可将其写为

$$\boldsymbol{x}=\frac{\boldsymbol{v}^*\boldsymbol{x}}{\boldsymbol{v}^*\boldsymbol{v}}\boldsymbol{v}+\boldsymbol{y}\triangleq\alpha\boldsymbol{v}+\boldsymbol{y}$$

其中，$\alpha\boldsymbol{v}\in\mathrm{span}\{\boldsymbol{v}\}$，$\boldsymbol{y}\in\mathrm{span}\{\boldsymbol{v}\}^{\perp}$；则

$$\boldsymbol{H}\boldsymbol{x}=\boldsymbol{x}-\frac{2}{\boldsymbol{v}^*\boldsymbol{v}}\boldsymbol{v}\boldsymbol{v}^*\boldsymbol{x}=\boldsymbol{x}-2\alpha\boldsymbol{v}=-\alpha\boldsymbol{v}+\boldsymbol{y}$$

即 $\boldsymbol{H}\boldsymbol{x}$ 与 $\boldsymbol{x}$ 在 $\mathrm{span}\{\boldsymbol{v}\}^{\perp}$ 方向有相同的分量，与在 $\boldsymbol{v}$ 方向的分量正好相差一个符号。也就是说，$\boldsymbol{H}\boldsymbol{x}$ 是 $\boldsymbol{x}$ 关于超平面 $\mathrm{span}\{\boldsymbol{v}\}^{\perp}$ 的镜面反射（见图 3.2.1）。因此，Householder 矩阵也称为反射矩阵。

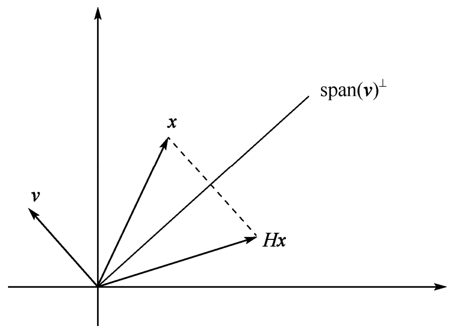

图 3.2.1　Householder 变换的几何意义

下面给出关于 Householder 矩阵的几个基本性质。

定理 3.2.3　假设 $\boldsymbol{H}\in C^{n\times n}$ 为一个 Householder 矩阵，那么

（1）$\boldsymbol{H}^*=\boldsymbol{H}$，即 $\boldsymbol{H}$ 是 Hermite 矩阵；

（2）$\boldsymbol{H}^*\boldsymbol{H}=\boldsymbol{I}$，即 $\boldsymbol{H}$ 是酉矩阵；

（3）$\boldsymbol{H}^2=\boldsymbol{I}$，即 $\boldsymbol{H}^{-1}=\boldsymbol{H}$；

（4）$\det(\boldsymbol{H})=-1$。

Householder 矩阵的一个非常重要的应用就是可以将一个向量除第一个元素以外的所有元素都化为零。我们首先给出引理 3.2.1。

引理 3.2.1　设 $\boldsymbol{x},\boldsymbol{y}\in C^n$ 为任意两个互异的向量，则存在一个 Householder 矩阵 $\boldsymbol{H}(\boldsymbol{v})$ 使得 $\boldsymbol{y}=\boldsymbol{H}(\boldsymbol{v})\boldsymbol{x}$ 的充要条件是 $\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2$ 且 $\boldsymbol{x}^*\boldsymbol{y}\in\mathbf{R}$。

证明　若 $\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2$ 且 $\boldsymbol{x}^*\boldsymbol{y}\in R$，则 $\boldsymbol{y}^*\boldsymbol{y}=\boldsymbol{x}^*\boldsymbol{x}$ 且 $\boldsymbol{x}^*\boldsymbol{y}=\boldsymbol{y}^*\boldsymbol{x}$。于是

$$\|\boldsymbol{x}-\boldsymbol{y}\|_2^2=(\boldsymbol{x}-\boldsymbol{y})^*(\boldsymbol{x}-\boldsymbol{y})=\boldsymbol{x}^*\boldsymbol{x}-\boldsymbol{y}^*\boldsymbol{x}-\boldsymbol{x}^*\boldsymbol{y}+\boldsymbol{y}^*\boldsymbol{y}=2(\boldsymbol{x}^*\boldsymbol{x}-\boldsymbol{y}^*\boldsymbol{x})$$

令 $\boldsymbol{v}=\boldsymbol{x}-\boldsymbol{y}$，则有

$$\boldsymbol{H}(\boldsymbol{v})\boldsymbol{x}=\boldsymbol{x}-\frac{2(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^*\boldsymbol{x}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}=\boldsymbol{x}-\frac{2(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}^*\boldsymbol{x}-\boldsymbol{y}^*\boldsymbol{x})}{2(\boldsymbol{x}^*\boldsymbol{x}-\boldsymbol{y}^*\boldsymbol{x})}=\boldsymbol{y}$$

即存在 Householder 矩阵 $\boldsymbol{H}(\boldsymbol{v})$ 使得 $\boldsymbol{y}=\boldsymbol{H}(\boldsymbol{v})\boldsymbol{x}$。

反之，如果存在 Householder 矩阵 $\boldsymbol{H}$ 使得 $\boldsymbol{y}=\boldsymbol{H}\boldsymbol{x}$，由于 $\boldsymbol{H}$ 是 Hermite 矩阵，所以 $\boldsymbol{x}^*\boldsymbol{y}=\boldsymbol{x}^*\boldsymbol{H}\boldsymbol{x}\in\mathbf{R}$。又因为 $\boldsymbol{H}$ 是酉矩阵，所以 $\|\boldsymbol{y}\|_2=\|\boldsymbol{H}\boldsymbol{x}\|_2=\|\boldsymbol{x}\|_2$。

由引理 3.2.1，我们可以得到定理 3.2.4。

定理 3.2.4　设 $\boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\in R^n$ 是一个非零向量，则存在 Householder 矩阵 $\boldsymbol{H}(\boldsymbol{v})$

使得 $\boldsymbol{H}(\boldsymbol{v})\boldsymbol{x}=\alpha\boldsymbol{e}_1$，其中，$\alpha=\|\boldsymbol{x}\|_2$（或$\alpha=-\|\boldsymbol{x}\|_2$），$\boldsymbol{e}_1=[1,0,\cdots,0]^{\mathrm{T}}\in R^n$。

设 $\boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\in R^n$ 是一个实的非零向量，下面讨论如何计算定理 3.2.2 中 Householder 矩阵 $\boldsymbol{H}(\boldsymbol{v})$ 对应的 Householder 向量 $\boldsymbol{v}$。由引理 3.2.1 的证明过程可知

$$\boldsymbol{v}=\boldsymbol{x}-\alpha\boldsymbol{e}_1=[x_1-\alpha,x_2,\cdots,x_n]^{\mathrm{T}}$$

在实际计算中，为了尽可能地减小舍入误差，我们通常避免将两个相近的数进行减法运算，否则就会损失有效数字。因此，通常取

$$\alpha=-\mathrm{sign}(x_1)\|\boldsymbol{x}\|_2$$

事实上，我们也可以取 $\alpha=\mathrm{sign}(x_1)\|\boldsymbol{x}\|_2$，但此时为了减小舍入误差，我们需要通过下面的公式来计算向量 $\boldsymbol{v}$ 的第一个分量 v_1

$$\alpha=\mathrm{sign}(x_1)\|\boldsymbol{x}\|_2,\quad v_1=x_1-\alpha=\frac{x_1^2-\|\boldsymbol{x}\|_2^2}{x_1+\alpha}=\frac{-\left(x_2^2+x_3^2+\cdots+x_n^2\right)}{x_1+\alpha}$$

$$v_1=\begin{cases} x_1-\alpha, & 若\mathrm{sign}(x_1)<0 \\ \dfrac{-\left(x_2^2+x_3^2+\cdots+x_n^2\right)}{x_1+\alpha}, & 其他 \end{cases}$$

无论怎样选取 α，我们都有 $\boldsymbol{H}=\boldsymbol{I}-\beta\boldsymbol{v}\boldsymbol{v}^*$ 其中

$$\beta=\frac{2}{\boldsymbol{v}^*\boldsymbol{v}}=\frac{2}{(x_1-\alpha)^2+x_2^2+\cdots+x_n^2}=\frac{2}{2\alpha^2-2\alpha x_1}=-\frac{1}{\alpha v_1}$$

在实数域中计算 Householder 向量 $\boldsymbol{v}$ 的算法如下，总运算量大约为 $3n$。

注　如果 $\boldsymbol{x}$ 是复向量，那么有没有相应的结论呢？如果有的话，其对应的 Householder 向量是什么呢？

在实际计算时，我们可以将向量 $\boldsymbol{v}$ 单位化，使得 $v_1=1$。这样，我们就不需要为 $\boldsymbol{v}$ 另外分配空间，而是将 $\boldsymbol{v}(2:n)$ 存放在 $\boldsymbol{x}(2:n)$ 中，因为经过 Householder 变换后，向量 $\boldsymbol{x}$ 除第一个分量外，其他分量都为零。同时，为了避免可能产生的溢出，我们也可以事先将 $\boldsymbol{x}$ 单位化，令 $\boldsymbol{x}=\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2}$。

算法 3.2.1　Householder 向量

```
function [beta, v] =House(x)
    n =length(x);
      sigma =dot(x(2:n), x(2:n));
```

```
    v =x;
    if  sigma ==0
          if   x(1) <0
               v(1) =2*x(1);
               beta =2/(v(1)*v(1));
          else
               v(1) =1;
               beta =0;
          end
    else
          alpha =sqrt(x(1)*x(1)+sigma);
          if   x(1) <0
               v(1) = x(1) -alpha;
          else
               v(1) =-sigma/(x(1)+alpha);
          end
               beta =2/(v(1)*v(1)+sigma);
    end
```

不难验证算法 3.2.1 具有很好的数值稳定性且满足$\|\tilde{\boldsymbol{H}}-\boldsymbol{H}\|_2=O(\varepsilon_u)$，其中，$\tilde{\boldsymbol{H}}$ 表示由上述算法计算得到的近似矩阵，ε_u 表示机器精度。

Householder 变换的运算量如下。

设 $\boldsymbol{A}\in R^{m\times n}$，$\boldsymbol{H}=\boldsymbol{I}-\beta\boldsymbol{v}\boldsymbol{v}^*\in R^m$，则

$$\boldsymbol{HA}=(\boldsymbol{I}-\beta\boldsymbol{v}\boldsymbol{v}^*)\boldsymbol{A}=\boldsymbol{A}-\beta\boldsymbol{v}(\boldsymbol{v}^*\boldsymbol{A})$$

因此，在进行 Householder 变换时，并不需要生成 Householder 矩阵，只需要 Householder 向量即可。上面矩阵相乘的总运算量大约为$4mn$。

3.2.4　Givens 变换

为简单起见，我们这里只讨论实数域中的 Givens 变换。设$\theta\in[0,2\pi]$，我们称矩阵

$$\boldsymbol{G}(i,j,\theta)=\begin{bmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & c & & s & & \\ & & & \ddots & & & \\ & & -s & & c & & \\ & & & & & \ddots & \\ & & & & & & 1\end{bmatrix}\in\mathbb{R}^{n\times n},\ i\leqslant j$$

为 Givens 变换（或 Givens 旋转，或 Givens 矩阵），其中，

$c=\cos\theta,\ \ s=\sin\theta$。即将单位矩阵的 (i,i) 和 (j,j) 位置上的元素用 c 代替，而 (i,j) 和 (j,i) 位置上的元素分别用 s 和 $-s$ 代替，得到的矩阵就是 $\boldsymbol{G}(i,j,\theta)$。

定理 3.2.5 $\boldsymbol{G}(i,j,\theta)$ 是正交矩阵，且 $\det[\boldsymbol{G}(i,j,\theta)]=1$。

值得注意的是当一个矩阵左乘一个 Givens 矩阵时，只会影响其第 i 行和第 j 行的元素。而当一个矩阵右乘一个 Givens 矩阵时，只会影响其第 i 列和第 j 列的元素。

算法 3.2.2 Givens 变换

```
% Given x=[a,b]^T ∈ ℝ^2 , compute c and s such that Gx=[r,0]^T where r=‖x‖_2
function    [c,s]=givens(a,b)
    if   b=0  then
       if   a≥0  then
                  c=1,s=0
       else
                  c=−1,s=0
       end if
            else
       if  |b|>|a|   then
                  τ=a/b, s=sign(b)/√(1+τ^2), c=sτ
       else
                  τ=b/a, c=sign(a)/√(1+τ^2), s=cτ
       end if
    end if
```

3.2.5 正交矩阵舍入误差分析

引理 3.2.2 设 $\boldsymbol{P}\in R^{n\times n}$ 是一个精确的 Householder 变换或 Givens 变换，$\boldsymbol{P}$ 是其浮点运算近似，则

$$\mathrm{fl}\left(\tilde{\boldsymbol{P}}\boldsymbol{A}\right)=\boldsymbol{P}(\boldsymbol{A}+\boldsymbol{E}),\ \mathrm{fl}\left(\boldsymbol{A}\tilde{\boldsymbol{P}}\right)=(\boldsymbol{A}+\boldsymbol{F})\boldsymbol{P}$$

其中

$$\|\boldsymbol{E}\|_2=O(\varepsilon_u)\|\boldsymbol{A}\|_2,\ \|\boldsymbol{F}\|_2=O(\varepsilon_u)\|\boldsymbol{A}\|_2$$

这说明一个矩阵进行 Householder 变换或 Givens 变换是向后稳定的。

定理 3.2.6　考虑对矩阵 $\boldsymbol{A}$ 进行一系列的正交变换，则有

$$\mathrm{fl}\left(\tilde{\boldsymbol{P}}\cdots\tilde{\boldsymbol{P}}\boldsymbol{A}\tilde{\boldsymbol{Q}}\cdots\tilde{\boldsymbol{Q}}\right)=\boldsymbol{P}_k\cdots\boldsymbol{P}_1\left(\boldsymbol{A}+\boldsymbol{E}\right)\boldsymbol{Q}_1\cdots\boldsymbol{Q}_k$$

其中，$\|\boldsymbol{E}\|_2=O\left(\varepsilon_u\right)\left(k\|\boldsymbol{A}\|_2\right)$。

定理 3.2.4　为了说明整个正交变换的计算过程是向后稳定的。不妨设 $\boldsymbol{X}$ 是一个非奇异的线性变换，$\tilde{\boldsymbol{X}}$ 表示其浮点运算近似。当 $\boldsymbol{X}$ 作用到 $\boldsymbol{A}$ 上时，我们有

$$\mathrm{fl}\left(\tilde{\boldsymbol{X}}\boldsymbol{A}\right)=\boldsymbol{X}\boldsymbol{A}+\boldsymbol{E}=\boldsymbol{X}\left(\boldsymbol{A}+\boldsymbol{X}^{-1}\boldsymbol{E}\right)\triangleq\boldsymbol{X}\left(\boldsymbol{A}+\boldsymbol{F}\right)$$

其中

$$\|\boldsymbol{E}\|_2=O(\varepsilon_u)\|\boldsymbol{X}\boldsymbol{A}\|_2\leqslant O(\varepsilon_u)\|\boldsymbol{X}\|_2\|\boldsymbol{A}\|_2$$

因此

$$\|\boldsymbol{F}\|_2=\left\|\boldsymbol{X}^{-1}\boldsymbol{E}\right\|_2\leqslant O\left(\varepsilon_u\right)\left\|\boldsymbol{X}^{-1}\right\|_2\|\boldsymbol{X}\|_2\|\boldsymbol{A}\|_2=O\left(\varepsilon_u\right)\kappa_2\left(\boldsymbol{X}\right)\|\boldsymbol{A}\|_2$$

此时，舍入误差将被放大 $\kappa_2(\boldsymbol{X})$ 倍。当 $\boldsymbol{X}$ 是正交变换时，$\kappa_2(\boldsymbol{X})$ 达到最小值 1，这就是在浮点运算中尽量使用正交变换的原因。

3.3　QR 分解

QR 分解是将一个矩阵分解为一个正交矩阵（酉矩阵）和一个三角矩阵的乘积。QR 分解被广泛应用于线性最小二乘问题的求解和矩阵特征值的计算。

3.3.1　QR 分解的存在性与唯一性

定理 3.3.1（QR 分解定理）　设 $A\in C^{m\times n}$（$m\geqslant n$）。则存在一个单位列正交矩阵 $\boldsymbol{Q}\in C^{m\times n}$（$\boldsymbol{Q}*\boldsymbol{Q}=\boldsymbol{I}_{n\times n}$）和一个上三角矩阵 $\boldsymbol{R}\in C^{n\times n}$，使得

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R} \tag{3.3.1}$$

若 $\boldsymbol{A}$ 列满秩，则存在一个具有正对角线元素的上三角矩阵 $\boldsymbol{R}$ 使得式（3.3.1）成立，且此时 QR 分解唯一，即 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 都唯一。

证明　设 $\boldsymbol{A}=[\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n]\in C^{m\times n}$。若 $\boldsymbol{A}$ 列满秩，即 $\mathrm{rank}(\boldsymbol{A})=n$，则 QR 分解（3.3.1）就是对 $\boldsymbol{A}$ 的列向量组进行 Gram-Schmidt 正交化过程的矩阵描述（见算法 3.3.1）。

算法 3.3.1　Gram-Schmidt 正交化

```
r_11 = ||a_1||_2
q_1 = a_1 / r_11
for j=2 to n do
    q_j = a_j
    for i=1 to j-1 do
        r_ij = q_i^* a_j   % q_i^* 表示共轭转置
        q_j = q_j - r_ij q_i
    end for
    r_jj = ||q_j||_2
    q_j = q_j / r_jj
end for
```

如果 $\boldsymbol{A}$ 不是列满秩的，我们可以进行类似的正交化过程：

- 如果 $\boldsymbol{a}_1=\boldsymbol{0}$，则令 $\boldsymbol{q}_1=\boldsymbol{0}$；否则令 $\boldsymbol{q}_1=\boldsymbol{a}_1/\|\boldsymbol{a}_1\|_2$。
- 对于 $j=2,3,\cdots,n$，计算 $\tilde{\boldsymbol{q}}_j=\boldsymbol{a}_j-\sum_{i=1}^{j-1}\left(\boldsymbol{q}_i^*\boldsymbol{a}_j\right)\boldsymbol{q}_i$。如果 $\tilde{\boldsymbol{q}}_j=\boldsymbol{0}$，则表明 $\boldsymbol{a}_j$ 可以由 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_{j-1}$ 线性表示，令 $\boldsymbol{q}_j=\boldsymbol{0}$ 否则令 $\boldsymbol{q}_j=\tilde{\boldsymbol{q}}_j/\|\tilde{\boldsymbol{q}}_j\|_2$。

这时可得到

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$$

其中，$\boldsymbol{Q}=[\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n]$ 列正交（但不是单位列正交），其列向量要么是单位向量，要么是零向量。这里的 $\boldsymbol{R}=[r_{ij}]_{n\times n}$ 是上三角矩阵，其中

$$r_{ij}=\begin{cases}\boldsymbol{q}_i^*\boldsymbol{a}_j, & i\leqslant j\\ 0, & i>j\end{cases}$$

如果 $\boldsymbol{Q}$ 的某一列 $\boldsymbol{q}_k=\boldsymbol{0}$，则 $\boldsymbol{R}$ 中对应的第 k 行就全为 0。

设 $\operatorname{rank}(\boldsymbol{A})=l<n$，则 $\boldsymbol{Q}$ 有 l 个非零列，设为 $\boldsymbol{q}_{i_1},\boldsymbol{q}_{i_2},\cdots,\boldsymbol{q}_{i_l}$。它们形成 C^m 中的一个单位正交向量，所以我们可以将其扩展成 C^m 中的一组标准正交基，即

$$\boldsymbol{q}_{i_1},\boldsymbol{q}_{i_2},\cdots,\boldsymbol{q}_{i_l},\tilde{\boldsymbol{q}}_1,\cdots,\tilde{\boldsymbol{q}}_{m-l}$$

然后我们用 $\tilde{\boldsymbol{q}}_1$ 替换 $\boldsymbol{Q}$ 中的第一个零列，用 $\tilde{\boldsymbol{q}}_2$ 替换 $\boldsymbol{Q}$ 中的第二个零列，依次类推，将 $\boldsymbol{Q}$ 中的所有零列都替换掉。将最后得到的矩阵记为 $\boldsymbol{Q}$，则 $\tilde{\boldsymbol{Q}}\in C^{m\times n}$ 单位列正交，且 $\tilde{\boldsymbol{Q}}\boldsymbol{R}=\boldsymbol{Q}\boldsymbol{R}$。由于 $\tilde{\boldsymbol{Q}}$ 中新添加的列向量正好与 $\boldsymbol{R}$ 中的零行相对应，所以我们有 QR 分解

$$\boldsymbol{A}=\tilde{\boldsymbol{Q}}\boldsymbol{R}$$

下面证明满秩矩阵 QR 分解的存在性和唯一性。

证明　存在性。由于 $\boldsymbol{A}$ 列满秩，由 Gram-Schmidt 正交化过程（算法 3.3.1）可知，存在上三角矩阵 $\boldsymbol{R}=[r_{ij}]_{n\times n}$ 满足 $r_{jj}>0$，使得 $\boldsymbol{A}=\boldsymbol{QR}$，其中，$\boldsymbol{Q}$ 单位列正交。

唯一性。假设 $\boldsymbol{A}$ 存在 QR 分解

$$\boldsymbol{A}=\boldsymbol{Q}_1\boldsymbol{R}_1=\boldsymbol{Q}_2\boldsymbol{R}_2$$

其中，$\boldsymbol{Q}_1,\boldsymbol{Q}_2\in C^{m\times n}$ 单位列正交；$\boldsymbol{R}_1,\boldsymbol{R}_2\in C^{n\times n}$ 为具有正对角元素的上三角矩阵，则有

$$\boldsymbol{Q}_1=\boldsymbol{Q}_2\boldsymbol{R}_2\boldsymbol{R}^{-1} \tag{3.3.2}$$

由于 $\boldsymbol{R}_1$、$\boldsymbol{R}_2$ 均为上三角矩阵，所以 $\boldsymbol{R}_2\boldsymbol{R}_1^{-1}$ 也是上三角矩阵，且其对角线元素为 $\boldsymbol{R}_2(i,i)/\boldsymbol{R}_1(i,i)$，$i=1,2,\cdots,n$。由式（3.3.2）可知

$$1=\|\boldsymbol{Q}_1\|_2=\left\|\boldsymbol{Q}_2\boldsymbol{R}_2\boldsymbol{R}_1^{-1}\right\|_2=\|\boldsymbol{R}_2\|_2$$

所以

$$\frac{\boldsymbol{R}_2(i,i)}{\boldsymbol{R}_1(i,i)}\leqslant 1,\quad i=1,2,\cdots,n$$

同理可证 $\boldsymbol{R}_1(i,i)/\boldsymbol{R}_2(i,i)\leqslant 1$。所以

$$\boldsymbol{R}_1(i,i)=\boldsymbol{R}_2(i,i),\quad i=1,2,\cdots,n$$

又 $\|\boldsymbol{Q}_1\|_{\mathrm{F}}=\mathrm{tr}(\boldsymbol{Q}_1\boldsymbol{Q}_1)=n$，所以由式（3.3.2）可知

$$\left\|\boldsymbol{R}_2\boldsymbol{R}_1^{-1}\right\|_{\mathrm{F}}^2=\left\|\boldsymbol{Q}_2\boldsymbol{R}_2\boldsymbol{R}_1^{-1}\right\|_{\mathrm{F}}^2=\|\boldsymbol{Q}_1\|_{\mathrm{F}}^2=n$$

由于 $\boldsymbol{R}_2\boldsymbol{R}_1^{-1}$ 的对角线元素都是 1，所以 $\boldsymbol{R}_2\boldsymbol{R}_1^{-1}$ 只能是单位矩阵，即 $\boldsymbol{R}_2=\boldsymbol{R}_1$。因此 $\boldsymbol{Q}_2=\boldsymbol{A}\boldsymbol{R}_2^{-1}=\boldsymbol{A}\boldsymbol{R}_1^{-1}=\boldsymbol{Q}_1$，即 $\boldsymbol{A}$ 的 QR 分解是唯一的。

由 QR 分解的存在性证明过程可知，当 $\boldsymbol{A}$ 不是满秩矩阵时，存在一个置换矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{AP}$ 的前 r 列是线性无关的，其中，$r=\mathrm{rank}(\boldsymbol{A})$。因此我们可以对 $\boldsymbol{AP}$ 进行 QR 分解，于是我们可以得到下面的结论。

推论 3.3.1　设 $A\in C^{m\times n}$（$m\geqslant n$）。则存在一个置换矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{AP}=\boldsymbol{Q}\begin{bmatrix}\boldsymbol{R}_{11} & \boldsymbol{R}_{12}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}_{n\times n}$$

其中，$\boldsymbol{Q}\in C^{m\times n}$ 单位列正交；$\boldsymbol{R}_{11}\in C^{r\times r}$ 是非奇异上三角矩阵。

在证明 QR 分解的存在性时，我们利用了 Gram-Schmidt 正交化过程。由于数值稳定性方面的原因，在实际计算中，我们一般不采用 Gram-Schmidt 正交化过程，而是采用修正的 Gram-Schmidt 过程（Modified Gram-Schmidt Process），即 MGS，具体过程如下所示。

算法 3.3.2　基于 MGS 的 QR 分解

```
%给定 A∈R^{m×n}，计算 Q=[q_1,…,q_n]∈R^{m×n} 且 R∈R^{n×n} 使得 A=QR
Set R=[r_ik]=0_{n×n}   (the n×n zero matrix)
if a_1=0 then
    q_1=0
else
    r_11=||a_1||_2
    q_1=a_1/||a_1||_2
end if
for k=2 to n do
    q_k=a_k
    for i=1 to k-1 do
        r_ik=q_i^T q_k
        q_k=q_k-r_ik q_i
    end for
    if q_k≠0 then
        r_kk=||q_k||_2
        q_k=q_k/r_kk
    end if
end for
```

注　由 MGS 得到的 QR 分解中，$\boldsymbol{Q}\in R^{m\times n}$，$\boldsymbol{R}\in R^{n\times n}$ 。

3.3.2　基于 Householder 变换的 QR 分解

由定理 3.2.4 可知，通过 Householder 变换，我们可以将任意一个非零变量 $\boldsymbol{x}\in R^n$ 转化成 $\|\boldsymbol{x}\|_2\boldsymbol{e}_1$，即除第一个元素外，其他元素都为零。下面我们就考虑通过 Householder 变换来实现矩阵的 QR 分解。

我们首先考虑 $m=n$ 时的情形。设矩阵 $\boldsymbol{A}\in R^{n\times n}$，令 $\boldsymbol{H}_1\in R^{n\times n}$ 为一个 Householder 变换，满足

$$\boldsymbol{H}_1\begin{bmatrix} a_{11}\\ a_{21}\\ \vdots\\ a_{n1}\end{bmatrix}=\begin{bmatrix} r_1\\ 0\\ \vdots\\ 0\end{bmatrix}$$

于是

$$
\boldsymbol{H}_1\boldsymbol{A}=\begin{bmatrix} r_1 & \tilde{a} & \cdots & \tilde{a} \\ 0 & & & \\ \vdots & & \tilde{\boldsymbol{A}}_2 & \\ 0 & & & \end{bmatrix}
$$

其中，$\tilde{\boldsymbol{A}}_2 \in R^{(n-1)\times(n-1)}$。同样地，我们可以构造一个 Householder 变换 $\boldsymbol{H}_2 \in R^{(n-1)\times(n-1)}$，将 $\tilde{\boldsymbol{A}}_2$ 的第一列中除第一个元素外的所有元素都化为 0，即

$$
\tilde{\boldsymbol{H}}\tilde{\boldsymbol{A}}=\begin{bmatrix} r_2 & \tilde{a} & \cdots & \tilde{a} \\ 0 & & & \\ \vdots & & \tilde{\boldsymbol{A}}_3 & \\ 0 & & & \end{bmatrix}
$$

令

$$
\boldsymbol{H}_2=\begin{bmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{H}}_2 \end{bmatrix}
$$

则

$$
\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A}=\begin{bmatrix} r_1 & \tilde{a} & \tilde{a} & \cdots & \tilde{a} \\ 0 & r_2 & \tilde{a} & \cdots & \tilde{a} \\ 0 & 0 & & & \\ \vdots & \vdots & & \tilde{\boldsymbol{A}}_3 & \\ 0 & 0 & & & \end{bmatrix}
$$

不断重复上述过程。这样，我们就得到一系列的矩阵

$$
\boldsymbol{H}_k=\begin{bmatrix} \boldsymbol{I}_{k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{H}} \end{bmatrix},\ k=1,2,\cdots,n-1
$$

使得

$$
\boldsymbol{H}_{n-1}\cdots\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A}=\begin{bmatrix} r_1 & \tilde{a} & \cdots & \tilde{a} \\ 0 & r_2 & \cdots & \tilde{a} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & r_n \end{bmatrix} \triangleq \boldsymbol{R}
$$

因为 Householder 变换都是正交矩阵，所以 $\boldsymbol{H}_1$、$\boldsymbol{H}_2$、$\cdots$、$\boldsymbol{H}_{n-1}$ 也都是正交矩阵。令

$$
\boldsymbol{Q}=(\boldsymbol{H}_{n-1}\cdots\boldsymbol{H}_2\boldsymbol{H}_1)^{-1}=\boldsymbol{H}_1^{-1}\boldsymbol{H}_2^{-1}\cdots\boldsymbol{H}_{n-1}^{-1}=\boldsymbol{H}_1\boldsymbol{H}_2\cdots\boldsymbol{H}_{n-1}
$$

则 $\boldsymbol{Q}$ 也是正交矩阵，且

$$\boldsymbol{A}=(\boldsymbol{H}_{n-1}\cdots\boldsymbol{H}_2\boldsymbol{H}_1)^{-1}\boldsymbol{R}=\boldsymbol{Q}\boldsymbol{R}$$

以上就是基于 Householder 变换的 QR 分解的具体实现过程。最后得到的上三角矩阵 $\boldsymbol{R}$ 就存放在 $\boldsymbol{A}$ 的上三角部分。矩阵 $\boldsymbol{Q}$ 可通过下面的算法实现

$$\begin{cases}\boldsymbol{Q}=\boldsymbol{I}_n\\ \boldsymbol{Q}=\boldsymbol{Q}\boldsymbol{H}_k, \quad k=1,2,\cdots,n-1\end{cases}$$

注　这里我们假定 $\boldsymbol{A}$ 是满秩的，如果 $\boldsymbol{A}$ 不是满秩矩阵，则可以考虑列主元 QR 分解。

如果 $m>n$，我们仍然可以通过上面的过程进行 QR 分解，只是最后我们会得到一个正交矩阵 $\boldsymbol{Q}\in R^{m\times m}$ 和一个上三角矩阵 $\boldsymbol{R}\in R^{m\times n}$，使得 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$。

如果不需要生成 $\boldsymbol{Q}$，则基于 Householder 变换的 QR 分解的总运算量大约为 $2mn^2-\dfrac{2}{3}n^3$。如果保留了每一步的 Householder 向量，则 $\boldsymbol{Q}$ 也可以通过下面的向后累积方法实现

$$\begin{aligned}&\boldsymbol{Q}=\boldsymbol{I}_n\\ &\boldsymbol{Q}=\boldsymbol{H}_k\boldsymbol{Q},\ k=n-1,n-2,\cdots,1\end{aligned}$$

这样做的好处是一开始 $\boldsymbol{Q}$ 会比较稀疏，随着迭代的进行，$\boldsymbol{Q}$ 才会慢慢变满。而前面的计算方法，第 1 步就将 $\boldsymbol{Q}$ 变成了一个满矩阵。采用这种方法计算 $\boldsymbol{Q}$ 的运算量大约为 $4(m^2n-mn^2+\dfrac{1}{3}n^3)$。

如果将 $\boldsymbol{Q}$ 写成下面的形式

$$\boldsymbol{Q}=\boldsymbol{I}+\boldsymbol{W}\boldsymbol{Y}^{\mathrm{T}}$$

则可以采用分块形式来计算 $\boldsymbol{W}$ 和 $\boldsymbol{Y}$，虽然运算量会稍有增长，但大多数运算是矩阵乘法，因此可以尽可能多地采用 3 级 BLAS 运算，如此效率可能会更高。

上面的算法只是关于利用 Householder 变换来实现 QR 分解的一个简单描述，并没有考虑运算量的问题。在实际计算中，我们通常会保留所有的 Householder 向量。由于第 k 步中 $\boldsymbol{H}_k$ 所对应的 Householder 向量 $\boldsymbol{v}_k$ 的长度为 $m-k+1$，所以我们先把 $\boldsymbol{v}_k$ 单位化，使得 $\boldsymbol{v}_k$ 的第一个元素为 1，这样就只需要存储 $\boldsymbol{v}_k(2:\mathrm{end})$，共 $m-k$ 个元素。所有的运算量大约为 $4m^2n-2mn^2+\dfrac{2}{3}n^3$，若 $m=n$，则运算量大约为 $\dfrac{8}{3}n^3$。

我们也可以考虑分块 Householder QR 分解，以便充分利用 3 级 BLAS 运算，提高计算效率。

3.3.3　列主元 QR 分解

当 $\boldsymbol{A}$ 不是满秩矩阵时，我们可以进行列主元 QR 分解。

定理 3.3.2（列主元 QR 分解）　设矩阵 $\boldsymbol{A}\in C^{m\times n}$（$m\geqslant n$），且 $\mathrm{rank}(\boldsymbol{A})=k<n$。则存在置换矩阵 $\boldsymbol{P}$ 和正交矩阵 $\boldsymbol{Q}\in C^{m\times m}$，使得

$$\boldsymbol{AP}=\boldsymbol{Q}\begin{bmatrix}\boldsymbol{R}_{11} & \boldsymbol{R}_{12}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}$$

其中，$\boldsymbol{R}_{11}\in C^{k\times k}$ 是非奇异上三角矩阵，且对角线元素满足 $r_{11}\geqslant r_{22}\geqslant\cdots\geqslant r_{kk}>0$。

列主元 QR 分解的实现过程与 QR 分解基本类似，只是在第 l 步时，需要选一个列主元，同时可能需要进行一个列交换。

假设经过 $l-1$ 步后，我们得到下面的分解

$$\boldsymbol{AP}^{(l-1)}=\boldsymbol{Q}^{(l-1)}\begin{bmatrix}\boldsymbol{R}_{11}^{(l-1)} & \boldsymbol{R}_{12}^{(l-1)}\\ \boldsymbol{0} & \boldsymbol{R}_{22}^{(l-1)}\end{bmatrix}\triangleq\boldsymbol{Q}^{(l-1)}\boldsymbol{R}^{(l-1)}，即 (\boldsymbol{Q}^{(l-1)})^{\mathrm{T}}\boldsymbol{AP}^{(l-1)}=\boldsymbol{R}^{(l-1)}$$

其中，$\boldsymbol{P}^{(l-1)}$ 是置换矩阵；$\boldsymbol{Q}^{(l-1)}$ 是正交矩阵；$\boldsymbol{R}^{(l-1)}\in R^{(l-1)\times(l-1)}$ 是非奇异上三角矩阵。

下面进行第 l 步：

（1）首先计算 $\boldsymbol{R}_{22}^{(l-1)}$ 的所有列的范数，如果范数都为 0，则 $\boldsymbol{R}_{22}^{(l-1)}=\boldsymbol{0}$，此时必有 $l-1=k$，算法结束。

（2）当 $l\leqslant k$ 时，$\boldsymbol{R}_{22}\neq\boldsymbol{0}$，记范数最大的列为第 i_l 列（如果有相等的，取其中一个即可）。若 $i_l\neq 1$，则交换 $\boldsymbol{R}^{(l-1)}$ 的第 l 列与第 i_l+l-1 列，并记相应的置换矩阵为 $\boldsymbol{P}_l$。

（3）以 $\boldsymbol{R}_{22}^{(l-1)}$ 的第 i_l 列为 Householder 向量，构造 Householder 变换 $\tilde{\boldsymbol{H}}_l$，并令 $\boldsymbol{H}_l=\begin{bmatrix}\boldsymbol{I}_{l-1} & \boldsymbol{0}\\ \boldsymbol{0} & \tilde{\boldsymbol{H}}_l\end{bmatrix}$，$\boldsymbol{P}^{(l)}=\boldsymbol{P}^{(l-1)}\boldsymbol{P}_l$，则

$$\boldsymbol{H}_l(\boldsymbol{Q}^{(l-1)})^{\mathrm{T}}\boldsymbol{AP}^{(l)}=\boldsymbol{H}_l\boldsymbol{R}^{(l-1)}\boldsymbol{P}_l=\begin{bmatrix}\boldsymbol{R}_{11}^{(l-1)} & \tilde{\boldsymbol{R}}\\ \boldsymbol{0} & \tilde{\boldsymbol{R}}\end{bmatrix}\triangleq\boldsymbol{R}^{(l)}$$

其中，$\tilde{\boldsymbol{R}}_{22}^{(l-1)}$ 的第一列除第一个元素外，其余元素都是零。且该元素的值等于 $\boldsymbol{R}_{22}^{(l-1)}$ 的第 i_l 列的范数。$\boldsymbol{Q}^{(l)}\triangleq\boldsymbol{Q}^{(l-1)}\boldsymbol{H}_l^{\mathrm{T}}$，则

$$\boldsymbol{AP}^{(l)}=\boldsymbol{Q}^{(l)}\boldsymbol{R}^{(l)}=\begin{bmatrix}\boldsymbol{R}_{11}^{(l)} & \boldsymbol{R}_{12}^{(l)}\\ \boldsymbol{0} & \boldsymbol{R}_{22}^{(l)}\end{bmatrix}$$

其中，$\boldsymbol{R}_{11}^{(l)} \in R^{l\times l}$ 为非奇异上三角矩阵。

依次类推，直到第 k 步，我们就可以得到 $\boldsymbol{A}$ 的列主元 QR 分解，其中，$\boldsymbol{R}_{11}$ 的对角线元素的递减关系是由选取列主元和 Householder 变换的性质造成的。

3.3.4　基于 Givens 变换的 QR 分解

我们同样可以利用 Givens 变换来做 QR 分解。

设矩阵 $\boldsymbol{A}\in R^{n\times n}$，首先构造一个 Givens 变换 $\boldsymbol{G}_{21}$，作用在矩阵 $\boldsymbol{A}$ 最前面的两行上，使

$$\boldsymbol{G}_{21}\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} \tilde{a}_{11} \\ 0 \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix}$$

由于 $\boldsymbol{G}_{21}$ 只改变矩阵的第 1 行和第 2 行的值，所以其他行保持不变。然后再构造一个 Givens 变换 $\boldsymbol{G}_{31}$，作用在 $\boldsymbol{G}_{21}\boldsymbol{A}$ 的第 1 行和第 3 行上，将其第 1 列的第 3 个元素化为零。由于 $\boldsymbol{G}_{31}$ 只改变矩阵的第 1 行和第 3 行的值，所以第 2 行的零元素保持不变。以此类推，我们可以构造一系列的 Givens 变换 $\boldsymbol{G}_{41}$、$\boldsymbol{G}_{51}$、…、$\boldsymbol{G}_{n1}$，使得 $\boldsymbol{G}_{n1}\cdots\boldsymbol{G}_{21}\boldsymbol{A}$ 的第 1 列中除第 1 个元素外，其他元素都化为零，即

$$\boldsymbol{G}_{n1}\cdots\boldsymbol{G}_{21}\boldsymbol{A}\begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix} = \begin{bmatrix} \tilde{a}_{11} \\ 0 \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix}$$

下面我们可以对第 2 列进行类似的处理。构造 Givens 变换 $\boldsymbol{G}_{32}$、$\boldsymbol{G}_{42}$、…、$\boldsymbol{G}_{n2}$，将第 2 列的第 3 至第 n 个元素全化为零，同时保持第 1 列不变。

以此类推，我们对其他列也进行类似的处理。最后，通过构造 $\dfrac{1}{2}n(n-1)$ 个 Givens 变换，可以将 $\boldsymbol{A}$ 转化成一个上三角矩阵 $\boldsymbol{R}$，即

$$\boldsymbol{R} = \boldsymbol{G}_{n,n-1}\cdots\boldsymbol{G}_{21}\boldsymbol{A}$$

令 $\boldsymbol{Q} = (\boldsymbol{G}_{n,n-1}\cdots\boldsymbol{G}_{21})^{\mathrm{T}}$。由于 Givens 变换是正交矩阵，所以 $\boldsymbol{Q}$ 也是正交矩阵。于是，我们就得到了矩阵 $\boldsymbol{A}$ 的 QR 分解

$$A = QR$$

与 Householder 变换一样，在进行 Givens 变换时，我们不需要显式地写出 Givens 矩阵。对于稠密矩阵而言，基于 Givens 变换的 QR 分解的运算量比 Householder 变换要多很多。

当需要连续应用一系列 Givens 变换时，我们可以使用快速 Givens 变换。即便如此，其速度仍然要慢于 Householder 变换。因此基于 Givens 变换的 QR 分解主要用于当矩阵的非零下三角元素相对较少的情形，如对上 Hessenberg 矩阵进行 QR 分解。

如果 $A \in R^{m\times n}$（$m > n$），我们仍然可以通过 Givens 变换进行 QR 分解。

下面是基于 Givens 变换的 QR 分解的算法描述。

算法 3.3.3　基于 Givens 变换的 QR 分解

```
%给定 A∈ℝ^{m×n}，计算 Q 和 R 使得 A=QR，Q∈ℝ^{m×m} 且 R∈ℝ^{m×n}
Set Q = I_{m×m}
for k=1 to n do
    for i=k+1 to m do
        [c, s] = givens(a_kk, a_ik)
        [A(k,k:n); A(i,k:n)] = G [A(k,k:n); A(i,k:n)]  where G = [c s; -s c]
        [Q(1:m,k), Q(1:m,i)] = [Q(1:m,k), Q(1:m,i)] G^T
    end for
end for
```

3.4　奇异值分解

奇异值分解（SVD）是一种非常实用的工具。

设矩阵 $A \in \mathbb{C}^{m\times n}$（$m \geqslant n$），则 $A^* A \in \mathbb{C}^{n\times n}$ 和 $AA^* \in \mathbb{C}^{m\times m}$ 都是 Hermite 半正定矩阵，且它们具有相同的非零特征值（都是正实数）。

定理 3.4.1（SVD 分解）　设矩阵 $A \in \mathbb{C}^{m\times n}$（$m \geqslant n$），则存在酉矩阵 $U \in \mathbb{C}^{m\times m}$ 和 $V \in \mathbb{C}^{n\times n}$ 使得

$$U^* A V = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \text{或 } A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^* \qquad (3.4.1)$$

其中，$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n) \in \mathbb{R}^{n\times n}$，且 $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$。式（3.4.1）称为矩阵 A 的奇异值分解，而 σ_1、σ_2、$\cdots$、σ_n 则称为矩阵 A 的奇异值。

证明 首先假设 $\boldsymbol{A}\neq\boldsymbol{0}$，否则只需要令 $\boldsymbol{\Sigma}=\boldsymbol{0}$ 即可，$\boldsymbol{U}$ 和 $\boldsymbol{V}$ 可以是任意酉矩阵。下面我们采用归纳法来证明定理 3.4.1 的结论。

当 $n=1$，$m\geqslant n$ 时，我们取 $\boldsymbol{\Sigma}=\|\boldsymbol{A}\|_2$，$\boldsymbol{V}=1$, $\boldsymbol{U}\in\mathbb{C}^{m\times m}$ 是第一列为 $\boldsymbol{u}_1=\boldsymbol{A}/\|\boldsymbol{A}\|_2$ 的酉矩阵，于是

$$\boldsymbol{A}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^*$$

为 $\boldsymbol{A}$ 的奇异值分解。

假设 $C^{(m-1)\times(n-1)}$ 中的矩阵都存在奇异值分解，下面证明 $\boldsymbol{A}\in C^{m\times n}$ 也存在奇异值分解。由 2 范数的定义可知，存在向量 $\boldsymbol{v}\in\mathbb{C}^n$ 满足 $\|\boldsymbol{v}\|_2=1$ 使得 $\|\boldsymbol{A}\|_2=\|\boldsymbol{A}\boldsymbol{v}\|_2$。令

$$\boldsymbol{u}=\frac{1}{\sigma}\boldsymbol{A}\boldsymbol{v}\in\mathbb{C}^m$$

其中，$\sigma=\|\boldsymbol{A}\|_2$，则 $\|\boldsymbol{u}\|_2=1$。我们将 $\boldsymbol{v}$ 和 $\boldsymbol{u}$ 都扩充成酉矩阵，即存在 $\tilde{\boldsymbol{U}}\in\mathbb{C}^{m\times(m-1)}$，$\tilde{\boldsymbol{V}}\in\mathbb{C}^{n\times(n-1)}$ 和 $[\boldsymbol{v},\tilde{\boldsymbol{V}}]\in\mathbb{C}^{n\times n}$ 都是酉矩阵。于是

$$\tilde{\boldsymbol{U}}^*\boldsymbol{A}\boldsymbol{v}=\tilde{\boldsymbol{U}}^*(\sigma\boldsymbol{u})=0,\ \ \boldsymbol{u}^*\boldsymbol{A}\boldsymbol{v}=\boldsymbol{u}^*(\sigma\boldsymbol{u})=\sigma$$

所以

$$\tilde{\boldsymbol{A}}=[\boldsymbol{u},\tilde{\boldsymbol{U}}]^*\boldsymbol{A}\left[\boldsymbol{v},\tilde{\boldsymbol{V}}\right]=\begin{bmatrix}\boldsymbol{u}^*\boldsymbol{A}\boldsymbol{v} & \boldsymbol{u}^*\boldsymbol{A}\tilde{\boldsymbol{V}}\\ \tilde{\boldsymbol{U}}\boldsymbol{A}\boldsymbol{v} & \tilde{\boldsymbol{U}}\boldsymbol{A}\tilde{\boldsymbol{V}}\end{bmatrix}=\begin{bmatrix}\sigma & \boldsymbol{u}^*\boldsymbol{A}\tilde{\boldsymbol{V}}\\ \boldsymbol{0} & \tilde{\boldsymbol{U}}\boldsymbol{A}\tilde{\boldsymbol{V}}\end{bmatrix}$$

又

$$\sigma=\|\boldsymbol{A}\|_2=\|\tilde{\boldsymbol{A}}\|_2=\|\tilde{\boldsymbol{A}}\|_2\geqslant\|\tilde{\boldsymbol{A}}\boldsymbol{e}_1\|_2=\|[\sigma,\boldsymbol{u}^*\boldsymbol{A}\tilde{\boldsymbol{V}}]^*\|_2=\sqrt{\sigma^2+\|\boldsymbol{u}^*\boldsymbol{A}\tilde{\boldsymbol{V}}\|_2^2}$$

所以 $\|\boldsymbol{u}^*\boldsymbol{A}\tilde{\boldsymbol{V}}\|_2=0$，即 $\boldsymbol{u}^*\boldsymbol{A}\tilde{\boldsymbol{V}}=\boldsymbol{0}$。于是

$$[\boldsymbol{u},\tilde{\boldsymbol{U}}]^*\boldsymbol{A}\left[\boldsymbol{v},\tilde{\boldsymbol{V}}\right]=\begin{bmatrix}\sigma & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}_1\end{bmatrix}$$

其中，$\boldsymbol{A}_1=\tilde{\boldsymbol{U}}\boldsymbol{A}\tilde{\boldsymbol{V}}\in\mathbb{C}^{(m-1)\times(n-1)}$。由归纳假设可知，$\boldsymbol{A}_1$ 存在奇异值分解，设为

$$\boldsymbol{A}_1=\boldsymbol{U}_1\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}_1^*$$

其中，$\boldsymbol{U}_1\in\mathbb{C}^{(m-1)\times(m-1)}$ 和 $\boldsymbol{V}_1\in\mathbb{C}^{(n-1)\times(n-1)}$ 都是酉矩阵；$\boldsymbol{\Sigma}_1\in\mathbb{R}^{(n-1)\times(n-1)}$ 是对角矩阵，且其对角线元素按降序排列。令

$$U=\left[u,\tilde{U}\right]\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1\end{bmatrix},\ V=\left[v,\tilde{V}\right]\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & V_1\end{bmatrix}$$

则$U\in C^{m\times m}$和$V\in C^{n\times n}$都是酉矩阵，且

$$\begin{aligned}U^*AV&=\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1\end{bmatrix}[u,\tilde{U}]^*A\left[v,\tilde{V}\right]\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & V_1\end{bmatrix}\\&=\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1\end{bmatrix}\begin{bmatrix}\sigma & \mathbf{0}\\ \mathbf{0} & A_1\end{bmatrix}\begin{bmatrix}1 & \mathbf{0}\\ \mathbf{0} & V_1\end{bmatrix}\\&=\begin{bmatrix}\sigma & \mathbf{0}\\ \mathbf{0} & U_1^*A_1V_1\end{bmatrix}=\begin{bmatrix}\Sigma\\ \mathbf{0}\end{bmatrix}\end{aligned}$$

其中，$\Sigma=\begin{bmatrix}\sigma & \mathbf{0}\\ \mathbf{0} & \Sigma_1\end{bmatrix}$。又

$$\sigma^2=\left\|U^*AV\right\|_2^2=\left\|\begin{bmatrix}\Sigma\\ \mathbf{0}\end{bmatrix}\right\|_2^2\Sigma=\rho\left(\begin{bmatrix}\sigma^2 & \mathbf{0}\\ \mathbf{0} & \Sigma_1^2\end{bmatrix}\right)$$

所以σ不小于Σ_1中的所有对角线元素，即Σ的对角线元素也按降序排列。因此，式（3.4.1）就是A的奇异值分解。由归纳法可知，定理 3.4.1 的结论成立。

下面介绍奇异值的一些基本性质。

定理 3.4.2　设$A=U\begin{bmatrix}\Sigma\\ \mathbf{0}\end{bmatrix}V^*$是$A\in\mathbb{C}^{m\times n}(m\geqslant n)$的奇异值分解，则下面结论成立：

（1）A^*A的特征值是σ_2，对应的特征向量是v_i，$i=1,2,\cdots,n$；

（2）AA的特征值是σ和$m-n$个零，对应的特征向量是u_i，$i=1,2,\cdots,m$；

（3）$\|A\|_2=\sigma_1$，$\|A\|_F=\sqrt{\sigma_1^2+\sigma_2^2++\sigma_n^2}$；

（4）若$\operatorname{rank}(A)=r\leqslant n$，则$\operatorname{Ran}(A)=\operatorname{span}\{u_1,u_2,\cdots,u_r\}$，$\operatorname{Ker}(A)=\operatorname{span}\{v_{r+1},v_{r+2},\cdots,v_n\}$；

（5）设$x\in\mathbb{C}^n$且$\|x\|_2=1$，则$\sigma^n\leqslant\|Ax\|_2\leqslant\sigma^1$；

（6）（酉不变性）设$X\in\mathbb{C}^{m\times m}$和$Y\in\mathbb{C}^{n\times n}$是酉矩阵，则$\sigma_i(X^*AY)=\sigma_i(A)$。

证明　留作练习。

定理 3.4.3　设$A=U\Sigma V^*$是$A\in\mathbb{C}^{m\times m}$的奇异值分解，则下面结论成立：

（1）$\left|\det(A)\right|=\sigma_1\sigma_2\cdots\sigma_n$；

（2）若A非奇异，则$\left\|A^{-1}\right\|_2=\sigma_n^{-1},\kappa_2(A)=\sigma_1\sigma_n$；

（3）若A是 Hermite 的，且$A=U\Lambda U^*$是A的特征值分解，即$U^*U=I$，$\Lambda=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，

其中，$|\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|$，则 $\boldsymbol{A}=\boldsymbol{U\Sigma V}$ 是 $\boldsymbol{A}$ 的奇异值分解，其中，$\sigma_i=|\lambda_i|$，$v_i=\operatorname{sign}(\lambda_i)u_i$，若 $\lambda_i=0$，则取 $\operatorname{sign}(\lambda_i)=1$；

（4）矩阵 $\boldsymbol{H}=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{A}^*\\ \boldsymbol{A} & \boldsymbol{0}\end{bmatrix}$ 的特征值是 $\pm\sigma_i$，对应的单位特征向量为 $\frac{1}{\sqrt{2}}\begin{bmatrix}\boldsymbol{v}_i\\ \pm\boldsymbol{u}_i\end{bmatrix}$。

证明　留作练习。

3.5　线性最小二乘问题的求解

本节我们讨论超定线性最小二乘问题的求解方法。

3.5.1　QR 分解法

这里假定 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$（$m\geqslant n$）是满秩矩阵。设 $\boldsymbol{A}$ 的 QR 分解为 $\boldsymbol{A}=\boldsymbol{QR}$，我们用 3 种不同的方法来推导线性最小二乘问题的解。

（1）将 $\boldsymbol{Q}$ 扩充成一个正交矩阵，记为 $[\boldsymbol{Q},\tilde{\boldsymbol{Q}}]\in\mathbb{R}^{m\times m}$。于是有

$$
\begin{aligned}
\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2 &= \left\|[\boldsymbol{Q},\tilde{\boldsymbol{Q}}]^{\mathrm{T}}(\boldsymbol{Ax}-\boldsymbol{b})\right\|_2^2\\
&= \left\|[\boldsymbol{Q},\tilde{\boldsymbol{Q}}]^{\mathrm{T}}(\boldsymbol{QRx}-\boldsymbol{b})\right\|_2^2\\
&= \left\|\begin{bmatrix}\boldsymbol{Rx}-\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}\\ -\tilde{\boldsymbol{Q}}\boldsymbol{b}\end{bmatrix}\right\|_2^2\\
&= \|\boldsymbol{Rx}-\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}\|_2^2+\|\tilde{\boldsymbol{Q}}\boldsymbol{b}\|_2^2\\
&\geqslant \|\tilde{\boldsymbol{Q}}\boldsymbol{b}\|_2^2
\end{aligned}
$$

等号成立当且仅当 $\boldsymbol{Rx}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$。所以最小二乘解为

$$\boldsymbol{x}_*=\boldsymbol{R}^{-1}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$$

（2）将 $\boldsymbol{b}$ 写成 $\boldsymbol{b}=\left(\boldsymbol{QQ}^{\mathrm{T}}+\boldsymbol{I}-\boldsymbol{QQ}^{\mathrm{T}}\right)\boldsymbol{b}$，则

$$
\begin{aligned}
\boldsymbol{Ax}-\boldsymbol{b} &= \boldsymbol{Ax}-\left(\boldsymbol{QQ}^{\mathrm{T}}+\boldsymbol{I}-\boldsymbol{QQ}^{\mathrm{T}}\right)\boldsymbol{b}\\
&= \left(\boldsymbol{Ax}-\boldsymbol{QQ}^{\mathrm{T}}\boldsymbol{b}\right)-\left(\boldsymbol{I}-\boldsymbol{QQ}^{\mathrm{T}}\right)\boldsymbol{b}
\end{aligned}
$$

由 $\boldsymbol{A}=\boldsymbol{QR}$ 和 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{I}$ 可知 $\boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{QQ}^{\mathrm{T}}\right)=\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{R}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{QQ}^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}-\boldsymbol{R}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{0}$，所以 $(\boldsymbol{Ax}-\boldsymbol{QQ}^{\mathrm{T}}\boldsymbol{b})^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{QQ}^{\mathrm{T}}\right)\boldsymbol{b}=-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{QQ}^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{QQ}^{\mathrm{T}}\right)\boldsymbol{b}=\boldsymbol{0}$，即 $\boldsymbol{Ax}-\boldsymbol{QQ}^{\mathrm{T}}\boldsymbol{b}$ 与 $(\boldsymbol{I}-\boldsymbol{QQ}^{\mathrm{T}})\boldsymbol{b}$ 正交。所以

$$\begin{aligned}\|Ax-b\|_2^2 &= \|Ax-QQ^{\mathrm{T}}b\|_2^2+\|(I-QQ^{\mathrm{T}})b\|_2^2\\ &= \|Rx-Q^{\mathrm{T}}b\|_2^2+\|(I-QQ^{\mathrm{T}})b\|_2^2\\ &\geqslant \|(I-QQ^{\mathrm{T}})b\|_2^2\end{aligned}$$

等号成立当且仅当 $\boldsymbol{Rx}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$。所以最小二乘解为

$$x_*=R^{-1}Q^{\mathrm{T}}b$$

事实上，$\boldsymbol{Q}$ 的列向量组成 Ran($\boldsymbol{A}$)的一组标准正交基，因此 $\boldsymbol{QQ}^{\mathrm{T}}$ 是 Ran($\boldsymbol{A}$)上的正交投影算子。由图 3.2.1 可知，x_* 满足

$$Ax=QQ^{\mathrm{T}}b$$

即 $QRx_*=QQ^{\mathrm{T}}b$，由此可得 $x_*=R^{-1}Q^{\mathrm{T}}b$。

（3）解正规方程。由定理 3.1.2 可知，最小二乘解为

$$x_*=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b=(R^{\mathrm{T}}Q^{\mathrm{T}}QR)^{-1}R^{\mathrm{T}}Q^{\mathrm{T}}b=(R^{\mathrm{T}}R)^{-1}R^{\mathrm{T}}Q^{\mathrm{T}}b=R^{-1}Q^{\mathrm{T}}b$$

用 QR 分解来求最小二乘解的运算量大约为 $2mn^2$（如果采用 Householder 变换的话，运算量大约为 $2mn^2-\frac{2}{3}n^3$）。当 $m\gg n$ 时，运算量大约为正规方程的 2 倍。当 m=n 时，两种算法的运算量几乎相同。通常 QR 分解法比较稳定，是求解最小二乘问题的首选方法，特别是当 $\boldsymbol{A}$ 的条件数较大（病态）时。

3.5.2　奇异值分解法

设 $A\in\mathbb{R}^{m\times n}$ 列满秩，$A=U\begin{bmatrix}\Sigma\\ 0\end{bmatrix}V^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的奇异值分解。令 U_n 为 $\boldsymbol{U}$ 的前 n 列组成的矩阵，即 $U=\begin{bmatrix}U_n & \tilde{U}\end{bmatrix}$，则

$$\begin{aligned}\|Ax-b\|_2^2 &= \left\|U\begin{bmatrix}\Sigma\\ 0\end{bmatrix}V^{\mathrm{T}}x-b\right\|_2^2\\ &= \left\|\begin{bmatrix}\Sigma\\ 0\end{bmatrix}V^{\mathrm{T}}x-[U_n\tilde{U}]^{\mathrm{T}}b\right\|_2^2\\ &= \left\|\begin{bmatrix}\Sigma V^{\mathrm{T}}x-U_n^{\mathrm{T}}b\\ -\tilde{U}b\end{bmatrix}\right\|_2^2\end{aligned}$$

$$
\begin{aligned}
&= \left\| \boldsymbol{\Sigma V}^{\mathrm{T}} \boldsymbol{x} - \boldsymbol{U}_n^{\mathrm{T}} \boldsymbol{b} \right\|_2^2 + \left\| \tilde{\boldsymbol{U}} \boldsymbol{b} \right\|_2^2 \\
&\geqslant \left\| \tilde{\boldsymbol{U}} \boldsymbol{b} \right\|_2^2
\end{aligned}
$$

等号成立当且仅当 $\boldsymbol{\Sigma V}^{\mathrm{T}} \boldsymbol{x} - \boldsymbol{U}^{\mathrm{T}} \boldsymbol{b} = \mathbf{0}$，即

$$
\boldsymbol{x} = \left(\boldsymbol{\Sigma V}^{\mathrm{T}} \right)^{-1} \boldsymbol{U}_n^{\mathrm{T}} \boldsymbol{b} = \boldsymbol{V \Sigma}^{-1} \boldsymbol{U}_n^{\mathrm{T}} \boldsymbol{b}
$$

为式（3.1.1）线性最小二乘问题的解。

3.6 最小二乘扰动分析

定理 3.6.1 设 $\boldsymbol{A} \in \mathbb{R}^{m \times n}$（$m \geqslant n$）且 $\operatorname{rank}(\boldsymbol{A}) = n$。设 $\boldsymbol{x}$ 是式（3.1.1）线性最小二乘问题的解，极小化 $\left\| (\boldsymbol{A} + \delta \boldsymbol{A}) \tilde{x} - (\boldsymbol{b} + \delta \boldsymbol{b}) \right\|_2$，则

$$
\frac{\left\| \tilde{\boldsymbol{x}} - \boldsymbol{x} \right\|_2}{\left\| \boldsymbol{x} \right\|_2} \leqslant \varepsilon \left[\frac{2\kappa_2(\boldsymbol{A})}{\cos\theta} + \kappa_2^2(\boldsymbol{A}) \tan\theta \right] + O\left(\varepsilon^2 \right)
$$

其中，$\kappa_2(\boldsymbol{A}) = \sigma_1(\boldsymbol{A}) / \sigma_n(\boldsymbol{A})$；$\theta$ 为 $\boldsymbol{b}$ 与 $\operatorname{Ran}(\boldsymbol{A})$ 的夹角；$\varepsilon \triangleq \max\left\{ \frac{\|\delta \boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}, \frac{\|\delta \boldsymbol{b}\|_2}{\|\boldsymbol{b}\|_2} \right\}$；并假定 $\varepsilon \kappa_2(\boldsymbol{A}) < 1$（确保 $\boldsymbol{A} + \delta \boldsymbol{A}$ 满秩，从而 $\tilde{\boldsymbol{x}}$ 唯一确定）。

我们记

$$
\kappa_{\mathrm{LS}} \triangleq \frac{2\kappa_2(\boldsymbol{A})}{\cos\theta} + \kappa_2^2(\boldsymbol{A}) \tan\theta
$$

这就是最小二乘问题的条件数。当 $\theta = 0$ 时，$\boldsymbol{b} \in \operatorname{Ran}(\boldsymbol{A}_2)$，此时 $\kappa_{\mathrm{LS}} = 2\kappa_2(\boldsymbol{A})$；当 $\theta = \pi/2$ 时，$\boldsymbol{b} \perp \operatorname{Ran}(\boldsymbol{A})$，此时最小二乘解为 $\boldsymbol{x} = \mathbf{0}$，而 $\kappa_{\mathrm{LS}} = \infty$；当 $0 < \theta < \pi/2$ 时，$\kappa_{\mathrm{LS}} = O\left[\kappa_2^2(\boldsymbol{A}) \right]$。

定义残量 $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A x}$, $\tilde{\boldsymbol{r}} = (\boldsymbol{b} + \delta \boldsymbol{b}) - (\boldsymbol{A} + \delta \boldsymbol{A}) \tilde{\boldsymbol{x}}$，我们有下面的性质

$$
\frac{\left\| \tilde{\boldsymbol{r}} - \boldsymbol{r} \right\|_2}{\left\| \boldsymbol{r} \right\|_2} \leqslant \varepsilon \left[1 + 2\kappa_2(\boldsymbol{A}) \right]
$$

当我们使用 QR 分解或 SVD 分解求解最小二乘问题时，由于采用的是正交变换，所以它们都是数值稳定的。而正规方程涉及求解方程组 $\boldsymbol{A}^{\mathrm{T}} \boldsymbol{A x} = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{b}$，其精度依赖于条件数 $\kappa_2\left(\boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} \right) = \kappa_2^2(\boldsymbol{A})$，因为其误差以 $\kappa_2^2(\boldsymbol{A})$ 的倍数增长。因此当 $\boldsymbol{A}$ 的条件数较大时，正规方程解的精度会大大降低。

3.7　习题

1．设 $A\in R^{m\times n}$，证明：$\mathrm{Rank}(A^{\mathrm{T}})=\mathrm{Rank}(A^{\mathrm{T}}A)$。

2．设 $B\in R^{m\times m}$ 是对称半正定矩阵，$A\in R^{m\times n}$，其中，$m\geqslant n$。试证明：矩阵

$$\begin{bmatrix} B & A \\ A^{\mathrm{T}} & \mathbf{0} \end{bmatrix}$$

非奇异的充要条件是 A 列满秩且矩阵 $[B\,A]$ 行满秩【A 列满秩且 $\mathrm{Ker}(B)\cap\mathrm{Ker}(A^{\mathrm{T}})=\{\mathbf{0}\}$】。

3．证明：设 $A\in R^{n\times n}$ 是正交矩阵，则 A 可表示成至多 n 个 Householder 变换的乘积。

4．证明：设 $A\in R^{n\times n}$ 是正交矩阵，则 A 可表示成至多 $\frac{1}{2}n(n-1)$ 个 Givens 变换的乘积。

5．设 $x=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\in R^n$ 是一个非零向量，H 是 Householder 矩阵，满足 $Hx=\alpha e_1$。证明：矩阵 H 的第一列与向量 x 平行。

6．设 $R\in \mathrm{R}^{n\times n}$ 是一个上三角矩阵，且对角线元素互不相同。证明：存在正交矩阵 Q，使得 $Q^{\mathrm{T}}RQ$ 为上三角矩阵，且对角线元素为 R 的对角线元素的降序排列。

7．（极分解）设 $A\in C^{n\times n}$。

证明：（1）存在酉矩阵 U 和唯一的 Hermite 半正定矩阵 P，使得 $A=PU$。

（2）若 A 非奇异，则 U 也是唯一的。

8．设 $A\in C^{n\times n}$。证明：A 可对角化当且仅当存在 Hermite 正定矩阵 P 使得 $P^{-1}AP$ 是正规矩阵。（提示：利用极分解，但不是对 A 进行极分解）

9．设 $A\in R^{n\times n}$，证明 $\rho(A)\leqslant\sigma_{\max}(A)$。