第五章　矩阵计算问题

在众多科学与工程学科，如物理、化学工程、统计学、经济学、生物学、信号处理、自动控制、系统理论、医学和军事工程等中，许多问题都可用数学建模成矩阵方程 $\boldsymbol{Ax} = \boldsymbol{b}$. 根据数据向量 $\boldsymbol{b} \in \mathbb{R}^{m\times 1}$ 和数据矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 的不同, 矩阵方程主要有以下三种类型:

(1) **适定方程组**：方程的个数与未知量的个数相等即 $m = n$，并且 $\boldsymbol{A}$ 满秩可逆，此时 $\boldsymbol{x}$ 有唯一的解。

(2) **超定方程组**：当上述 $m > n$ 时, 并且数据矩阵 $\boldsymbol{A}$ 和数据向量 $\boldsymbol{b}$ 均已知，其中之一或者二者可能存在误差或者干扰。

(3) **欠定方程组**：当上述 $m < n$ 时, 数据矩阵 $\boldsymbol{A}$ 和数据向量 $\boldsymbol{b}$ 均已知，但未知向量 $\boldsymbol{x}$ 可能要求为稀疏向量。

我们这里引进线性方程并给出它的标准形式 $\boldsymbol{Ax} = \boldsymbol{y}$，其中 $\boldsymbol{x} \in \mathbb{R}^n$ 是未知变量，$\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 是参数矩阵，$\boldsymbol{y} \in \mathbb{R}^m$ 是已知向量。线性方程构成了数值线性代数的基础，它们的解法是许多优化方法的关键。事实上，解线性方程组问题 $\boldsymbol{Ax} = \boldsymbol{y}$ 可以被看成**优化问题**，即关于 $\boldsymbol{x}$，最小化 $\|\boldsymbol{Ax} - \boldsymbol{y}\|^2$ 。我们描述线性方程组解得集合并且当线性方程组正确解不存在的情况下，讨论求解线性方程组近似解的方法。随后引出最小二乘问题以及它的变体、解的数值敏感性及其解决方法，它们与矩阵分解的关系 (例如 QR 分解和 SVD) 也将被介绍。

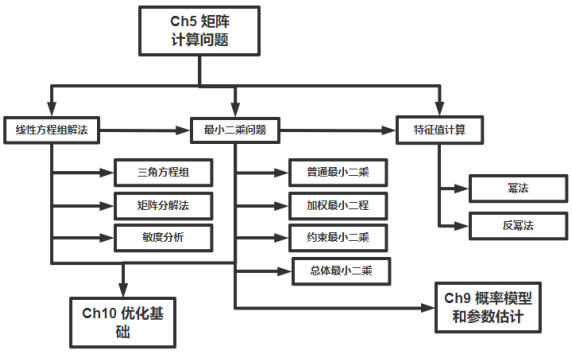

图 5.1: 本章导图

5.1 线性方程组的直接解法

5.1.1 线性方程组问题

在工程问题中，线性方程组描述了变量之间最基本的关系。线性方程在各个科学分支中无处不在，例如弹性力学、电阻网络、曲线拟合等。线性方程构成了线性代数的核心并时常作为优化问题的约束条件。由于许多优化算法的迭代过程非常依赖线性方程组的解，所以它也是许多优化算法的基础。下面我们以一个例子来说明，线性方程组如何解决上面的问题。

例 5.1.1. *(三点测距问题)* 三角测量是一种确定点位置的方法，给定距离到已知控制点 *(锚点)*，三边测量可以应用于许多不同的领域，如地理测绘、地震学、导航（例如 *GPS* 系统）等。在图*5.2*中，三个测距点 $a_1, a_2, a_3 \in \mathbb{R}^2$ 的坐标是已知的，并且从点 $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$ 到测距点的距离为 d_1, d_2, d_3，$\boldsymbol{x}$ 的未知坐标与距离测量有关，可以由下面非线性方程组描述

$$\|\boldsymbol{x} - \boldsymbol{a}_1\|_2^2 = d_1^2, \quad \|\boldsymbol{x} - \boldsymbol{a}_2\|_2^2 = d_2^2, \quad \|\boldsymbol{x} - \boldsymbol{a}_3\|_2^2 = d_3^2 \tag{5.1}$$

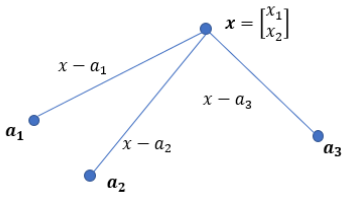

图 5.2: 三点测量位置图例。点 $\boldsymbol{x}$ 处，我们测量距三个测距点 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 的距离，以便确定 $\boldsymbol{x}$ 的坐标。

通过第一个方程减去另外两个方程，我们获得了两个 $\boldsymbol{x}$ 的线性方程组。

$$2\left(\boldsymbol{a}_2 - \boldsymbol{a}_1\right)^{\mathrm{T}} \boldsymbol{x} = d_1^2 - d_2^2 + \|\boldsymbol{a}_2\|_2^2 - \|\boldsymbol{a}_1\|_2^2$$
$$2\left(\boldsymbol{a}_3 - \boldsymbol{a}_1\right)^{\mathrm{T}} \boldsymbol{x} = d_1^2 - d_3^2 + \|\boldsymbol{a}_3\|_2^2 - \|\boldsymbol{a}_1\|_2^2$$

也就是说，原始非线性方程组(5.1)的每个解也可以看作线性方程组的解。使用方程组标准形式 $\boldsymbol{Ax} = \boldsymbol{y}$*(标准形式的定义在下一小节给出)* 可以描述为：

$$A = \begin{bmatrix} 2\left(\boldsymbol{a}_2 - \boldsymbol{a}_1\right)^{\mathrm{T}} \\ 2\left(\boldsymbol{a}_3 - \boldsymbol{a}_1\right)^{\mathrm{T}} \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} d_1^2 - d_2^2 + \|\boldsymbol{a}_2\|_2^2 - \|\boldsymbol{a}_1\|_2^2 \\ d_1^2 - d_3^2 + \|\boldsymbol{a}_3\|_2^2 - \|\boldsymbol{a}_1\|_2^2 \end{bmatrix} \tag{5.2}$$

上述问题的解，将在后面详细讨论。

我们先回顾线性方程组中的一般概念。

含 n 个未知量 x_1,x_2,...,x_n，m 个方程的线性方程组的一般形式为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{5.3}$$

若记

$$\boldsymbol{A} = (a_{ij})^{m\times n} \text{，} \boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \text{，} \boldsymbol{b} = (b_1, b_2, \cdots, b_m)^{\mathrm{T}}, \tag{5.4}$$

则方程组(5.3)可表为如下的矩阵形式：

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \tag{5.5}$$

当 $\boldsymbol{b} = \boldsymbol{0}$ 时，方程组(5.5)所对应的齐次线性方程组为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}. \tag{5.6}$$

定义 5.1.1. 矩阵 $\boldsymbol{A} = (a_{ij})^{m\times n}$, 称为方程组(5.5) 的 **系数矩阵**，而矩阵 $\tilde{\boldsymbol{A}} = (\boldsymbol{A}, \boldsymbol{b})$ 称为它的 **增广矩阵**。方程组(5.6)称为方程组(5.5) 的 **导出组**。

定义 5.1.2. 给定方程组

$$\bar{\boldsymbol{A}}\boldsymbol{x} = \bar{\boldsymbol{b}}. \tag{5.7}$$

其中 $\bar{\boldsymbol{A}} = (\bar{a}_{ij})^{m\times n}$，$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，$\bar{\boldsymbol{b}} = (\bar{b}_1, \bar{b}_2, \cdots, \bar{b}_m)^{\mathrm{T}}$。当 $\boldsymbol{x}^0 = (x_1^0, x_2^0, \cdots x_n^0)^{\mathrm{T}}$ 是方程组(5.7) 的解向量时，若它也是方程组(5.5) 的解向量，则称方程组(5.5) 与方程组(5.7)是 **同解方程组**。

定理 5.1.1. 对方程组(5.5)的系数矩阵 $\boldsymbol{A}$ 及右端作相同的行初等变换，所得到的新方程组与原方程组同解。

定义 5.1.3. 设 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_t$ 是齐次线性方程组(5.6)的解向量组，如果 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_t$ 线性无关，且方程组(5.6)的任意解向量 $\boldsymbol{\eta}$ 都可由 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_t$ 线性表出，则称解向量组 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_t$ 为方程组(5.6)的一个 **基础解系**。

定理 5.1.2. 设齐次线性方程组(5.6)的系数矩阵 $\boldsymbol{A}$ 的秩为 r，此时

(1) 方程组(5.6)有非零解的必要充分条件是 $r < n$。

(2) 若 $r < n$，则方程组(5.6)一定有基础解系。基础解系不是唯一的，但任两个基础解系必等价，且每一个基础解系所含解向量的个数都等于 $n - r$。

(3) 若 $r < n$，设 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$ 是方程组(5.6) 的一个基础解系，则它的一般解为

$$\boldsymbol{\eta} = \lambda_1\boldsymbol{\eta}_1 + \lambda_2\boldsymbol{\eta}_2 + \cdots + \lambda_{n-r}\boldsymbol{\eta}_{n-r}, \tag{5.8}$$

其中 $\lambda_i (i = 1, 2, \cdots, n-r)$ 是数域 $\mathbb{K}$ 中的任意常数.

定理 5.1.3. 方程组(5.5)有解的必要充分条件是: $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\tilde{\boldsymbol{A}})$. 矩阵 $\tilde{\boldsymbol{A}}$ 为它的 **增广矩阵**.

定理 5.1.4. 设 $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\tilde{\boldsymbol{A}}) = r$,$\boldsymbol{\gamma}_0$ 是非齐次方程组(5.5)的一个解向量 *(常称为* **特解***)*, $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$ 是其导出组(5.6)的一个基础解系, 则方程组(5.5)的解向量均可表为:

$$\boldsymbol{\gamma} = \boldsymbol{\gamma}_0 + \boldsymbol{\eta} = \boldsymbol{\gamma}_0 + \lambda_1\boldsymbol{\eta}_1 + \lambda_2\boldsymbol{\eta}_2 + \cdots + \lambda_{n-r}\boldsymbol{\eta}_{n-r},$$

其中 $\lambda_i (i = 1, 2, \cdots, n-r)$ 是数域 $\mathbb{K}$ 中的任意常数（这种形式的解向量常称为一般解）.

对增广矩阵 $\tilde{\boldsymbol{A}}$ 进行初等行变换将其化为阶梯型矩阵，写出相应的阶梯型方程组.

（1）若 $r = n$, 则阶梯型方程组形如:

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n = d_1, \\ \qquad\quad c_{22}x_2 + \cdots + c_{2n}x_n = d_2, \\ \qquad\qquad\qquad\qquad \cdots\cdots\cdots \\ \qquad\qquad\qquad\qquad\quad c_{nn}x_n = d_n. \end{cases} \tag{5.9}$$

其中 $c_{ii} \neq 0\,(i = 1, 2, \cdots, n)$. 依次由第 n 个,第 $n-1$ 个,$\cdots$,第一个方程组可解得 $x_n, x_{n-1}, \cdots, x_1$, 由此即得方程组(5.3) 的唯一解 $x_1, x_2, \cdots, x_n$.

（2）若 $r < n$, 则阶梯型方程组可表为:

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1r}x_r = d_1 - c_{1,r+1}x_{r+1} - \cdots - c_{1n}x_n, \\ \qquad\quad c_{22}x_2 + \cdots + c_{2r}x_r = d_2 - c_{2,r+1}x_{r+1} - \cdots - c_{2n}x_n, \\ \qquad\qquad\qquad\qquad\qquad\qquad \cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad\quad c_{rr}x_r = d_r - c_{r,r+1}x_{r+1} - \cdots - c_{rn}x_n. \end{cases} \tag{5.10}$$

其中 $c_{ii} \neq 0\,(i = 1, 2, \cdots, r)$. 此时方程组(5.3) 有无穷多组解. 若令 $x_{r+1} = x_{r+2} = \cdots = x_n = 0$, 则可由方程组(5.10) 求得一个特解 $\boldsymbol{\gamma}_0 = (\delta_1, \delta_2, \cdots, \delta_r, 0, 0, \cdots, 0)$，再由其导出组的一个基础解系 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$，可得方程组(5.3)的一般解.

线性方程组的解集被定义为:

$$S \doteq \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}\} \tag{5.11}$$

用 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n \in \mathbb{R}^n$ 表示矩阵 $\boldsymbol{A}$ 的列, 即 $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n]$. $\boldsymbol{A}\boldsymbol{x}$ 仅仅表示矩阵 $\boldsymbol{A}$ 的列与向量 $\boldsymbol{x}$ 中各个元素的加权和:

$$\boldsymbol{A}\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n \tag{5.12}$$

通过定义，我们能够看出, 无论 $\boldsymbol{x}$ 的值是什么，$\boldsymbol{A}\boldsymbol{x}$ 生成了由矩阵 $\boldsymbol{A}$ 的列张成的子空间. 向量 $\boldsymbol{A}\boldsymbol{x} \in \mathrm{Range}\,(\boldsymbol{A})$. 若 $\boldsymbol{y} \notin \mathrm{Range}\,(\boldsymbol{A})$，则线性方程组没有解. 因此解集 $\mathbb{S}$ 为空. 等价地，线性方程组有解当且仅当 $\boldsymbol{y} \in \mathrm{Range}\,(\boldsymbol{A})$.

回顾定义，设 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$.$\boldsymbol{A}$ 的值域定义为

$$\mathrm{Range}\,(\boldsymbol{A}) = \{\boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}^n\} \tag{5.13}$$

易证 Range $(\boldsymbol{A}) = \mathbf{span}\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$，其中 $\boldsymbol{a}_i$ 为 $\boldsymbol{A}$ 的列向量. $\boldsymbol{A}$ 的零空间定义为:

$$\text{Null}(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = 0\} \tag{5.14}$$

它的维数记为 nullity $(\boldsymbol{A})$

一个子空间 $\mathbb{S} \subset \mathbb{R}^n$ 的正交补定义为:

$$\mathbb{S}^{\perp} = \{\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} = 0, \forall \boldsymbol{x} \in \mathbb{S}\} \tag{5.15}$$

从矩阵值域的角度，给出定理5.1.3的证明：

定理 5.1.5. *方程组的解(5.5)存在的充分必要条件是 $rank(\boldsymbol{A}) = rank([\boldsymbol{A}, \boldsymbol{b}])$.*

证明. 必要性：设存在 $\boldsymbol{x}$ 使 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$，则 $\boldsymbol{b}$ 是 $\boldsymbol{A}$ 的列向量的线性组合，即 $\boldsymbol{b} \in \mathcal{R}(\boldsymbol{A})$。这说明 $\mathcal{R}([\boldsymbol{A}, \boldsymbol{b}]) = \mathcal{R}(\boldsymbol{A})$，所以有 $rank(\boldsymbol{A}) = rank([\boldsymbol{A}, \boldsymbol{b}])$。

充分性：若 $rank(\boldsymbol{A}) = rank([\boldsymbol{A}, \boldsymbol{b}])$ 成立，则 $\boldsymbol{b} \in \mathcal{R}(\boldsymbol{A})$，即 $\boldsymbol{b}$ 可表示为 $\boldsymbol{b} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{a}_i$，这里 $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]$，所以令 $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^T$，即有 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$。 □

定理 5.1.6. *假定方程组(5.5)的解存在, 并且假定 $\boldsymbol{x}$ 是其任一给定的解, 则(5.5)全部解的集合是*

$$\boldsymbol{x} + Null(\boldsymbol{A}) \tag{5.16}$$

证明. 如果 $\boldsymbol{y}$ 满足(5.5), 则 $\boldsymbol{A}(\boldsymbol{y} - \boldsymbol{x}) = 0$, 即 $(\boldsymbol{y} - \boldsymbol{x}) \in \text{Null}(\boldsymbol{A})$, 于是有 $\boldsymbol{y} = \boldsymbol{x} + (\boldsymbol{y} - \boldsymbol{x}) \in \boldsymbol{x} + \text{Null}(\boldsymbol{A})$. 反之，如果 $\boldsymbol{y} \in x + \text{Null}(\boldsymbol{A})$, 则存在 $\boldsymbol{z} \in \text{Null}(\boldsymbol{A})$, 使 $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$, 从而有 $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. □

定理5.1.6告诉我们, 只要知道了方程组(5.5)的一个解, 便可以用它及 Null $(\boldsymbol{A})$ 中向量的和得到(5.5)的全部解. 由此可知, 方程组(5.5)的解唯一, 只有当 Null $(\boldsymbol{A})$ 中仅有零向量才行.

推论 5.1.1. *方程组(5.5)的解唯一的充分必要条件是 $nullity(\boldsymbol{A}) = \mathbf{0}$.*

对于不同类型的方程组，解的数量情况有各自的特点。

超定方程组

线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 中线性方程的个数大于未知变量的个数时，我们说 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 是超定的. 或者说矩阵 $\boldsymbol{A}^{m \times n}$ 的行数大于列数: $m > n$. 假设 $\boldsymbol{A}$ 是一个列满秩矩阵, 也就是说 rank$(\boldsymbol{A}) = n$, 则我们可以得出 Null $(\boldsymbol{A}) = 0$.

因此线性方程组的解要么没有解，要么有唯一解. 在超定方程组中，$\boldsymbol{y} \notin$ Range $(\boldsymbol{A})$ 是很常见的，因此引入近似解的概念，近似解使得 $\boldsymbol{A}\boldsymbol{x}$ 与 $\boldsymbol{y}$ 在合适的度量下距离最小.

欠定方程组

线性方程组 $\boldsymbol{Ax} = \boldsymbol{y}$ 中未知变量的个数小于方程组的个数时, 我们说线性方程组是欠定的. 或者说 $\boldsymbol{A}^{m\times n}$ 的列数大于行数:$m < n$. 假设 $\boldsymbol{A}$ 是一个行满秩矩阵, 也就是说 $\text{rank}(\boldsymbol{A}) = m$, $\text{Range}\,(\boldsymbol{A}) = \mathbb{R}^m$. 则根据定理5.1.6:

$$\text{rank}(\boldsymbol{A}) + \dim(\text{Null}\,(\boldsymbol{A})) = n \tag{5.17}$$

因此 $\dim(\text{Null}\,(\boldsymbol{A})) = n - m > 0$. 此时线性方程组有解且有无限多个解，并且解集的维度是 $n - m$。在所有可能的解中，我们总是对具有最小范数的解很感兴趣.(后面详细讨论).

适定方程组

线性方程组 $\boldsymbol{Ax} = \boldsymbol{y}$ 中线性方程的个数等于未知变量的个数时，我们说 $\boldsymbol{Ax} = \boldsymbol{y}$ 是适定方程组. 或者说 $\boldsymbol{A}^{m\times n}$ 的列数等于行数:$m = n$. 如果系数矩阵是满秩的即 $\boldsymbol{A}$ 可逆, $\boldsymbol{A}^{-1}$ 唯一且有 $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$. 在这种情况下，线性方程组的解是唯一的:

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y} \tag{5.18}$$

注意，实际中我们几乎不会通过先求 $\boldsymbol{A}^{-1}$ 再乘以向量 $\boldsymbol{y}$ 的方式求解 $\boldsymbol{x}$. 而是通过数值方法 (比如之前学过的 LU 分解，Cholesky 分解) 来计算线性方程组非奇异方程组的解.

5.1.2 三角形线性方程组

在前面，我们考虑了理论上如何对一个一般的线性方程组求解。

接下来，我们将考虑如何使用计算机对一个线性方程组求解，尤其是一个规模巨大的线性方程组。

首先来考虑一个稍微简单些的情况，三角形线性方程组。

定义 5.1.4. 如果一个矩阵 $\boldsymbol{A}$ 主对角线以上所有元素为 0, 则称其为**下三角矩阵**。

如果一个矩阵 $\boldsymbol{A}$ 主对角线以下的所以元素为 0, 则称其为**上三角矩阵**。

定义 5.1.5. 如果一个线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 的系数矩阵 $\boldsymbol{A}$ 是上三角形矩阵或者下三角矩阵，我们则称其为**上三角形线性方程组**或者**下三角形线性方程组**。

针对上三角形线性方程组和下三角形线性方程组，我们可以分别用两种特别的方法解出方程组的解。

接下来，我们分别介绍这两种方法。

我们利用**前代法**计算下三角形线性方程组。

注意，我们要求系数矩阵主对角线上元素均非 0。从而保证方程组有且仅有一个解。

$$\begin{pmatrix} a_{11} & & & & \\ a_{21} & a_{22} & & & \\ a_{31} & a_{32} & a_{33} & & \\ \vdots & \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

其中 $a_{11}, a_{22}, ..., a_{nn}$ 非 0.

在前代法的第 (k) 个循环中，我们将会遇到下面这样一个形式

$$\begin{pmatrix} a_{11} & & & & & & & \\ 0 & a_{22} & & & & & & \\ 0 & 0 & a_{33} & & & & & \\ \vdots & \vdots & \vdots & \ddots & & & & \\ 0 & 0 & 0 & \dots & a_{kk} & & & \\ 0 & 0 & 0 & \dots & a_{k+1k} & a_{k+1k+1} & & \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \dots & a_{nk} & a_{nk+1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_k^{(k-1)} \\ b_{k+1}^{(k-1)} \\ \vdots \\ b_n^{(k-1)} \end{pmatrix}$$

此时我们将第 k 列从第 $k+1$ 行到第 n 行化为 0，同时更新 b。

$$\begin{pmatrix} a_{11} & & & & & & & \\ 0 & a_{22} & & & & & & \\ 0 & 0 & a_{33} & & & & & \\ \vdots & \vdots & \vdots & \ddots & & & & \\ 0 & 0 & 0 & \dots & a_{kk} & & & \\ 0 & 0 & 0 & \dots & 0 & a_{k+1k+1} & & \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \dots & 0 & a_{nk+1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2^{(1)} \\ b_3^{(2)} \\ \vdots \\ b_k^{(k-1)} \\ b_{k+1}^{(k)} \\ \vdots \\ b_n^{(k)} \end{pmatrix}$$

所以前代法，就从前 (x_1) 往后 (x_n) 来依次求解。

回代法则恰好相反，他是从后往前一次求解。

回代法是用于上三角形的线性方程组求解。

同样我们要求其系数矩阵对角线上元素非 0.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

与前代法类似，在回代法第 $(n-k+1)$ 个循环内。

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k-1} & a_{1k} & 0 & \dots & 0 \\ & a_{22} & \dots & a_{2k-1} & a_{2k} & 0 & \dots & 0 \\ & & \ddots & & \vdots & \vdots & & \vdots \\ & & & a_{k-1k-1} & a_{k-1k} & 0 & \dots & 0 \\ & & & & a_{kk} & 0 & \dots & 0 \\ & & & & & a_{k+1k+1} & \dots & 0 \\ & & & & & & \ddots & \vdots \\ & & & & & & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(n-k)} \\ b_2^{(n-k)} \\ \vdots \\ b_{k-1}^{(n-k)} \\ b_k^{(n-k)} \\ b_{k+1}^{(n-k-1)} \\ \vdots \\ b_n \end{pmatrix}$$

此时我们将第 k 列从第 1 行到第 $k-1$ 行化为 0，同时更新 b。

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k-1} & 0 & 0 & \dots & 0 \\ & a_{22} & \dots & a_{2k-1} & 0 & 0 & \dots & 0 \\ & & \ddots & & \vdots & \vdots & & \vdots \\ & & & a_{k-1k-1} & 0 & 0 & \dots & 0 \\ & & & & a_{kk} & 0 & \dots & 0 \\ & & & & & a_{k+1k+1} & \dots & 0 \\ & & & & & & \ddots & \vdots \\ & & & & & & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(n-k+1)} \\ b_2^{(n-k+1)} \\ \vdots \\ b_{k-1}^{(n-k+1)} \\ b_k^{(n-k)} \\ b_{k+1}^{(n-k-1)} \\ \vdots \\ b_n \end{pmatrix}$$

5.1.3 矩阵分解解线性方程组

LU 分解解方程

如果线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的系数矩阵 $\boldsymbol{A}$ 可以进行 LU 分解，则我们可以对 $\boldsymbol{A}$ 先进行 LU 分解.

令 $\boldsymbol{A}=\boldsymbol{LU}$, 我们则可以得到方程组 $\boldsymbol{LUx}=\boldsymbol{b}$。令 $\boldsymbol{Ux}=\boldsymbol{y}$，则我们可以得到两个线性方程组。

$$\boldsymbol{Ly}=\boldsymbol{b}$$

$$\boldsymbol{Ux}=\boldsymbol{y}$$

然后利用前代法求出 $\boldsymbol{y}$, 再利用回代法求解出原线性方程组的解 $\boldsymbol{x}$。

因为在计算机上求解方程，我们还需要考虑资源问题，为了节约资源，下面给出一种紧凑的求解方式。

给定矩阵 $\boldsymbol{A}$ 和向量 $\boldsymbol{b}$，我们先对 $\boldsymbol{A}$ 进行 LU 分解。并且使用 $\boldsymbol{A}$ 的上三角部分存储上三角矩阵，用下三角部分存储下三角矩阵。

比如矩阵

$$\begin{pmatrix} 3 & 2 & -1 \\ 6 & 6 & -2 \\ -3 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

就可以使用

$$\begin{pmatrix} 3 & 2 & -1 \\ 2 & 2 & 0 \\ -1 & 2 & -1 \end{pmatrix}$$

来存储。

然后我们来重新给出 LU 分解的算法流程。

$u_{1i} = a_{1i}, \quad i = 1, \cdots, n;$

$l_{i1} = a_{i1}/u_{11}, i = 2, \cdots, n;$

对 $k = 2$,$\cdots$,n, 计算

$u_{ki} = a_{ki} - \sum_{r=1}^{k-1} l_{kr}u_{ri}, \quad i = k, \cdots, n;$

$l_{ik} = (a_{ik} - \sum_{r=1}^{k-1} l_{kr}u_{ri})/u_{kk}, \quad k = 2, \cdots, n.$ 最后再使用前代法和回代法求出最终的解。

例 5.1.2. 求解 $\begin{pmatrix} 3 & 2 & -1 \\ 6 & 6 & -2 \\ -3 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -5 \end{pmatrix}$

我们先对 $\begin{pmatrix} 3 & 2 & -1 \\ 6 & 6 & -2 \\ -3 & 2 & 0 \end{pmatrix}$ LU 分解。

$$\to \begin{pmatrix} 3 & 2 & -1 \\ 2 & 6 & -2 \\ -1 & 2 & 0 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & -1 \\ 2 & 2 & 0 \\ -1 & 4 & -1 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & -1 \\ 2 & 2 & 0 \\ -1 & 2 & -1 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & -1 \\ 2 & 2 & 0 \\ -1 & 2 & -1 \end{pmatrix}$$

然后再进行前代法，注意此处，我们只关心 $\boldsymbol{y}$，不关注 $\boldsymbol{L}$ 如何变换，故 $\boldsymbol{L}$ 并不需要实际上去化为 0。可得

$$\hat{\boldsymbol{y}} = \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix}$$

最后进行回代法，便得解 $(1, -1, 1)^{\mathrm{T}}$.

其他类型的分解解法

对于正定矩阵还可以使用 Cholesky 分解来进行求解。对于方程组

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

其中 $\boldsymbol{A}$ 为正定矩阵。

我们可将其分解得到 $\boldsymbol{A} = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}$。然后再按照上述前代法和回代法的方式得到方程的解。

对于非方阵的情况，我们还可以使用 QR 分解来求解方程组。

对于方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$，

我们先对 $\boldsymbol{A}$ 进行 QR 分解，得到 $\boldsymbol{Q}\boldsymbol{R}\boldsymbol{x} = \boldsymbol{b}$ 即可得一上三角形线性方程组的同解方程组 $\boldsymbol{R}\boldsymbol{x} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$。再利用回代法求解。

5.1.4 敏度分析与其他方法

在本节中，我们分析了数据中小扰动对非奇异方阵线性方程解的影响。

输入的扰动敏感性

令 $\boldsymbol{x}$ 为线性方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 的解，其中 $\boldsymbol{A}$ 为非奇异方阵，且 $\boldsymbol{y} \neq 0$。假设我们通过向它添加一个小的扰动项 $\delta\boldsymbol{y}$ 来略微改变 $\boldsymbol{y}$，并将 $\boldsymbol{x} + \delta\boldsymbol{x}$ 称为扰动方程组的解：

$$\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y} + \delta\boldsymbol{y}$$

我们的关键问题是：如果 $\delta\boldsymbol{y}$ 变小，$\delta\boldsymbol{x}$ 将会不会变小？我们从上面的公式看出，并且从 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 的事实看，扰动 $\delta\boldsymbol{x}$ 本身就是线性方程组的解。

$$\boldsymbol{A}\delta\boldsymbol{x} = \delta\boldsymbol{y}$$

并且，由于认为 $\boldsymbol{A}$ 是可逆的，我们可以写成

$$\delta\boldsymbol{x} = \boldsymbol{A}^{-1}\delta\boldsymbol{y}$$

采用该方程两边的欧几里德范数得出

$$\|\delta\boldsymbol{x}\|_2 = \left\|\boldsymbol{A}^{-1}\delta\boldsymbol{y}\right\|_2 \leq \left\|\boldsymbol{A}^{-1}\right\|_2 \|\delta\boldsymbol{y}\|_2$$

其中 $\left\|\boldsymbol{A}^{-1}\right\|_2$ 是 $\boldsymbol{A}^{-1}$ 的谱 (最大奇异值) 范数。类似地，从 $\boldsymbol{Ax}=\boldsymbol{y}$ 得出 $\|\boldsymbol{y}\|_2=\|\boldsymbol{Ax}\|_2\leq\|\boldsymbol{A}\|_2\|\boldsymbol{x}\|_2$，因此

$$\|\boldsymbol{x}\|_2^{-1}\leq\frac{\|\boldsymbol{A}\|_2}{\|\boldsymbol{y}\|_2}$$

将上面两个公式相乘，我们得到

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}\leq\left\|\boldsymbol{A}^{-1}\right\|_2\|\boldsymbol{A}\|_2\frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2}$$

这个结果是我们正在寻找的，因为它将“输入项”$\boldsymbol{y}$ 的相对变化与“输出”$\boldsymbol{x}$ 的相对变化联系起来。数量

$$\kappa(\boldsymbol{A})=\left\|\boldsymbol{A}^{-1}\right\|_2\|\boldsymbol{A}\|_2,1\leq\kappa(\boldsymbol{A})\leq\infty$$

是矩阵 $\boldsymbol{A}$ 的条件数，见方程 (5.5)。大的 $\kappa(\boldsymbol{A})$ 意味着 $\boldsymbol{y}$ 上的扰动在 $\boldsymbol{x}$ 上被极大地放大，即方程对输入数据的变化非常敏感。如果 $\boldsymbol{A}$ 是奇异的，那么 $\kappa=\infty$。非常大的 $\kappa(\boldsymbol{A})$ 表明 $\boldsymbol{A}$ 接近奇异；我们说在这种情况下 $\boldsymbol{A}$ 是病态的。我们在以下引理中总结了我们的发现。

引理 5.1.1. *(对于输入的敏感性) 令 $\boldsymbol{A}$ 为非奇异方阵，$\boldsymbol{x},\delta\boldsymbol{x}$ 满足*

$$\boldsymbol{Ax}=\boldsymbol{y}$$
$$\boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{y}+\delta\boldsymbol{y}$$

然后它认为

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}\leq\kappa(\boldsymbol{A})\frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2}$$

其中 $\kappa(\boldsymbol{A})=\left\|\boldsymbol{A}^{-1}\right\|_2\|\boldsymbol{A}\|_2$ 是矩阵 $\boldsymbol{A}$ 的条件数

系数矩阵中的扰动敏感性

接下来我们考虑 $\boldsymbol{A}$ 矩阵的扰动对 $\boldsymbol{X}$ 的影响。令 $\boldsymbol{Ax}=\boldsymbol{y}$ 并且令 $\delta\boldsymbol{A}$ 为一个扰动，满足下面等式

$$(\boldsymbol{A}+\delta\boldsymbol{A})(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{y},\quad\text{对于一些}\delta\boldsymbol{x}$$

那么有

$$\boldsymbol{A}\delta\boldsymbol{x}=-\delta\boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x})$$

因此 $\delta\boldsymbol{x}=-\boldsymbol{A}^{-1}\delta\boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x})$。则

$$\|\delta\boldsymbol{x}\|_2=\left\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x})\right\|_2\leq\left\|\boldsymbol{A}^{-1}\right\|_2\|\delta\boldsymbol{A}\|_2\|\boldsymbol{x}+\delta\boldsymbol{x}\|_2$$

并且

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x}+\delta\boldsymbol{x}\|_2}\leq\left\|\boldsymbol{A}^{-1}\right\|_2\|\boldsymbol{A}\|_2\frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}$$

我们再次看到只有在条件数不是太大时，小扰动 $\frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}\ll1$ 对 $\boldsymbol{x}$ 的相对影响才很小。就是说，它离 1 不太远，$\kappa(\boldsymbol{A})\simeq1$。这个会在下一个引理中总结。

引理 5.1.2. *(系数矩阵中的扰动敏感性) 令 $\boldsymbol{A}$ 为非奇异方阵，$\boldsymbol{x}, \delta\boldsymbol{A}, \delta\boldsymbol{x}$ 满足*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$$

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y}$$

然后它认为

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} \leq \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}$$

对 A,y 联合扰动的敏感性

我们最后考虑了 $\boldsymbol{A}$ 和 $\boldsymbol{y}$ 的同时扰动对 $\boldsymbol{x}$ 的影响。令 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$，并且令 $\delta\boldsymbol{A}, \delta\boldsymbol{y}$ 为扰动，满足下面等式

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y} + \delta\boldsymbol{y}, \text{对于一些} \delta\boldsymbol{x}$$

然后，$\boldsymbol{A}\delta\boldsymbol{x} = \delta\boldsymbol{y} - \delta\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x})$，因此 $\delta\boldsymbol{x} = \boldsymbol{A}^{-1}\delta\boldsymbol{y} - \boldsymbol{A}^{-1}\delta\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x})$。则

$$\begin{aligned}
\|\delta\boldsymbol{x}\|_2 &= \left\|\boldsymbol{A}^{-1}\delta\boldsymbol{y} - \boldsymbol{A}^{-1}\delta\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x})\right\|_2 \\
&\leq \left\|\boldsymbol{A}^{-1}\delta\boldsymbol{y}\right\|_2 + \left\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x})\right\|_2 \\
&\leq \left\|\boldsymbol{A}^{-1}\right\|_2 \|\delta\boldsymbol{y}\|_2 + \left\|\boldsymbol{A}^{-1}\right\| \delta\boldsymbol{A} \|_2 \| \boldsymbol{x} + \delta\boldsymbol{x} \|_2
\end{aligned}$$

接着，上式除以 $\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2$，

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} \leq \left\|\boldsymbol{A}^{-1}\right\|_2 \frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2} \frac{\|\boldsymbol{y}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} + \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}$$

但是 $\|\boldsymbol{y}\|_2 = \|\boldsymbol{A}\boldsymbol{x}\|_2 \leq \|\boldsymbol{A}\|_2 \|\boldsymbol{x}\|_2$，因此

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} \leq \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2} \frac{\|\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} + \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}$$

下一步，我们根据 $\|\boldsymbol{x}\|_2 = \|\boldsymbol{x} + \delta\boldsymbol{x} - \delta\boldsymbol{x}\|_2 \leq \|\boldsymbol{x} + \delta\boldsymbol{x}\|_2 + \|\delta\boldsymbol{x}\|_2$ 去写

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} \leq \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2} \frac{\|\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} + \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}$$

从中我们得到

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} \leq \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2} \left(1 + \frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2}\right) + \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}$$

因此

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} \leq \frac{\kappa(\boldsymbol{A})}{1 - \kappa(\boldsymbol{A}) \frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2}} \left(\frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2} + \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}\right)$$

扰动的“放大因子”是受 $\frac{\kappa(\boldsymbol{A})}{1 - \boldsymbol{x}(\boldsymbol{A}) \frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2}}$ 的约束。因此，该界限小于某些给定的 γ，如果

$$\kappa(\boldsymbol{A}) \leq \frac{\gamma}{1 + \gamma \frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2}}$$

因此，我们看到关节扰动的影响仍然由 $\boldsymbol{A}$ 的条件数控制，如下所述

引理 5.1.3. *(对 A,y 扰动的敏感性) 令 $\boldsymbol{A}$ 为非奇异方阵, 令 $\gamma > 1$ 已知，并且令 $\boldsymbol{x}, \delta\boldsymbol{y}, \delta\boldsymbol{A}, \delta\boldsymbol{x}$ 满足下面等式*

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$$

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y} + \delta\boldsymbol{y}$$

然后

$$\kappa(\boldsymbol{A}) \leq \frac{\gamma}{1 + \gamma\frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2}}$$

这意味着

$$\frac{\|\delta\boldsymbol{x}\|_2}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|_2} \leq \gamma\left(\frac{\|\delta\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2} + \frac{\|\delta\boldsymbol{A}\|_2}{\|\boldsymbol{A}\|_2}\right)$$

5.2　最小二乘问题

本节的重点是如何通过解决最小二乘问题来解决超定方程组和欠定方程组。在本节开始我们先看一个在解决最小二乘问题中常用的定理:

定理 5.2.1. *下面 2 条性质成立:*

1. $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 是一个列满秩矩阵 (即 $\text{rank}(\boldsymbol{A}) = n$) 当且仅当 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是可逆的.

2. $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 是一个行满秩矩阵 (即 $\text{rank}(\boldsymbol{A}) = m$) 当且仅当 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是可逆的.

证明. 对于第一点. 如果 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 不是可逆的, 则存在 $\boldsymbol{x} \neq 0$ 使得 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = 0$.　$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = 0$, 因此 $\boldsymbol{A}\boldsymbol{x} = 0$. 所以 $\boldsymbol{A}$ 不是一个列满秩矩阵. 反之, 如果 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是可逆的, 对于每个 $\boldsymbol{x} \neq 0$,$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \neq 0$, 也能推出对于每一个非零的 $\boldsymbol{x}$,$\boldsymbol{A}\boldsymbol{x} \neq 0$. 第二点的证明过程与第一点的证明过程相似. □

5.2.1　最小二乘问题与线性回归

最小二乘问题多产生于线性回归或者数据拟合问题。例如，假定给出 m 个点 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_m$ 和这 m 个点上的试验或观测数据 $y_1, \cdots, y_m$，其中 $t_i \in \mathbb{R}^n, i = 1, \cdots, m$。再假定给出在 $\boldsymbol{t}_i$ 上取值的 n 个已知函数 $a_1(t), \cdots, a_n(t)$。考虑 a_i 的线性组合

$$f(x;t) = x_1a_1(t) + x_2a_2(t) + \cdots + x_na_n(t)$$

我们希望在 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_m$ 点上 $f(x;t)$ 能最佳地逼近这些数据 $y_1, \cdots, y_m$。为此，若定义残量

$$r_i(x) = y_i - \sum_{j=1}^{n} x_j a_j(t_i), \quad i = 1, \ldots, m$$

则问题成为：估计参数 $x_1, \cdots, x_n$，使残量 $r_1, \cdots, r_m$ 尽可能小。可用矩阵向量形式表示为

$$\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$$

其中

$$\boldsymbol{A}=\begin{bmatrix} a_1(t_1) & \cdots & a_n(t_1) \\ \vdots & & \vdots \\ a_1(t_m) & \cdots & a_n(t_m) \end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

$$\boldsymbol{x}=(x_1,\cdots,x_n)^T,\quad \boldsymbol{r}(\boldsymbol{x})=(\boldsymbol{r}_1(\boldsymbol{x}),\cdots,\boldsymbol{r}_m(\boldsymbol{x}))$$

当 $m=n$ 时，我们可以要求 $\boldsymbol{r}(\boldsymbol{x})=0$，则估计 x 的问题就可用上一节关于线性方程组中讨论的方法解决。当 $m>n$ 时，一般不可能使所有残量为 0，但我们可要求残差向量在某种范数意义下达最小，最小二乘问题就是求 $\boldsymbol{x}$ 使得残差向量 $\boldsymbol{r}(\boldsymbol{x})$ 在 2 范数意义下最小。下面给出最小二乘问题的定义。

定义 5.2.1. 给定矩阵 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 和向量 $\boldsymbol{b}\in\mathbb{R}^m$，确定 $x\in\mathbb{R}^n$ 使得

$$\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2=\|\boldsymbol{r}(\boldsymbol{x})\|_2=\min_{\boldsymbol{y}\in\mathbb{R}^n}\|\boldsymbol{r}(\boldsymbol{y})\|_2=\min_{\boldsymbol{y}\in\mathbb{R}^n}\|\boldsymbol{A}\boldsymbol{y}-\boldsymbol{b}\|_2 \tag{5.19}$$

该问题称为最小二乘问题，简记为 *LS(Least Squares)* 问题，其中 $\boldsymbol{r}(\boldsymbol{x})$ 称为残差向量。

如果残差向量 $\boldsymbol{r}$ 线性依赖于 $\boldsymbol{x}$，则称其为线性最小二乘问题；如果 $\boldsymbol{r}$ 非线性的依赖于 $\boldsymbol{x}$，则称其为非线性最小二乘问题。我们主要讨论线性最小二乘问题，简称最小二乘问题。对于残差向量选择不同的范数，便得到不同的问题，我们主要讨论残差向量选择 2 范数的情况。

最小二乘问题的解 $\boldsymbol{x}$ 又称为线性方程组(5.20)的最小二乘解。

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\quad \boldsymbol{A}\in\mathbb{R}^{m\times n} \tag{5.20}$$

即残差向量 $\boldsymbol{r}(\boldsymbol{x})$ 的 2 范数最小的意义下满足方程组(5.20)。

根据 m 与 n 以及矩阵 $\boldsymbol{A}$ 的秩 $r(\boldsymbol{A})$ 的不同，最小二乘问题可分为下面几种情况：

(1)$m=n$

$(1a)r(\boldsymbol{A})=m=n$

$(1b)r(\boldsymbol{A})<m=n$

(2)$m>n$

$(2a)r(\boldsymbol{A})=n<m$

$(2b)r(\boldsymbol{A})<n<m$

(3)$m<n$

$(3a)r(\boldsymbol{A})=m<n$

$(3b)r(\boldsymbol{A})<m<n$

其中 $m>n$ 对应的方程组称为超定方程组，$m<n$ 对应的方程组称为欠定方程组。不同的情况通常需要不同的方法去处理，我们主要讨论 (2a) 情况的最小二乘问题的性质与求解。

定理 5.2.2. 线性最小二乘问题的解总是存在的，而且其解唯一的充分必要条件是 $null(\boldsymbol{A}) = 0$。

证明. 因为 $\mathbb{R}^m = \mathcal{R}(\boldsymbol{A}) \oplus \mathcal{R}(\boldsymbol{A})^{\perp}$，所以向量 $\boldsymbol{b}$ 可以唯一地表示 $\boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2$，其中 $\boldsymbol{b}_1 \in \mathcal{R}(\boldsymbol{A}), \boldsymbol{b}_2 \in \mathcal{R}(\boldsymbol{A})^{\perp}$。于是对任意 $\boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{b}_1 - \boldsymbol{A}\boldsymbol{x} \in \mathcal{R}(\boldsymbol{A})$ 且与 $\boldsymbol{b}_2$ 正交，从而

$$\begin{aligned} \|\boldsymbol{r}(\boldsymbol{x})\|_2^2 &= \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2 = \|(\boldsymbol{b}_1 - \boldsymbol{A}\boldsymbol{x}) + \boldsymbol{b}_2\|_2^2 \\ &= \|\boldsymbol{b}_1 - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \|\boldsymbol{b}_2\|^2 \end{aligned}$$

由此可得，$\|\boldsymbol{r}(\boldsymbol{x})\|_2^2$ 达到极小当且仅当 $\|\boldsymbol{b}_1 - \boldsymbol{A}\boldsymbol{x}\|_2^2$ 达到极小，而 $\boldsymbol{b}_1 \in \mathcal{R}(\boldsymbol{A})$ 又蕴含着 $\|\boldsymbol{b}_1 - \boldsymbol{A}\boldsymbol{x}\|_2^2$ 达到极小地充要条件 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}_1$，综上，由 $\boldsymbol{b}_1 \in \mathcal{R}(\boldsymbol{A})$ 和推论 5.1.1 推出定理的结论成立。 □

记最小二乘问题的解集为 $\mathbb{S}_{LS}$，即

$$\mathbb{S}_{LS} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x}\text{是 LS 问题的解}\}$$

则由定理5.2.2知，$\mathbb{S}_{LS}$ 总是非空的，而且它仅有一个元素的充要条件是 $\boldsymbol{A}$ 的列线性无关。

5.2.2　欠定问题的最小范数解

我们要考虑下一个情况：当矩阵 $\boldsymbol{A}$ 的列数比行数多：$m < n$. 假设矩阵 $\boldsymbol{A}$ 是行满秩，我们有 $\dim\{\text{Null}(\boldsymbol{A})\} = n - m > 0$, 因此得出 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ 有无数个解并且解的集合是 $\mathbb{S}_{\boldsymbol{x}} = \{\boldsymbol{x} : \boldsymbol{x} = \tilde{\boldsymbol{x}} + \boldsymbol{z}, \boldsymbol{z} \in \text{Null}(\boldsymbol{A})\}$, 其中 $\tilde{\boldsymbol{x}}$ 是任意满足 $\boldsymbol{A}\tilde{\boldsymbol{x}} = \boldsymbol{y}$ 的向量。

(我们有兴趣从这个解的集合 $\mathbb{S}_{\boldsymbol{x}}$ 中挑选出一个满足最小欧几里得范数的解 $\boldsymbol{x}^*$)

我们想解决的问题:

$$\min_{\boldsymbol{x}:\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}} \|\boldsymbol{x}\|_2$$

这个式子等价于 $\min_{\boldsymbol{x}\in\mathbb{S}_{\boldsymbol{x}}} \|\boldsymbol{x}\|_2$.（唯一的）解 $\boldsymbol{x}^*$ 必须与 $\text{Null}(\boldsymbol{A})$ 相互垂直，等价地，$\boldsymbol{x}^* \in \text{Range}\left(\boldsymbol{A}^T\right)$ 这个意味着存在 ζ，$\boldsymbol{x}^* = \boldsymbol{A}^T\zeta$. 因为 $\boldsymbol{x}^*$ 是方程组的解，必须满足 $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{y}$, 就等于说 $\boldsymbol{A}\boldsymbol{A}^T\zeta = \boldsymbol{y}$.

因为矩阵 $\boldsymbol{A}$ 是行满秩，$\boldsymbol{A}\boldsymbol{A}^T$ 是可逆的并且有唯一的 ζ 是方程组的解。$\zeta = (\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{y}$.

这样我们得到了唯一的最小范数解：

$$\boldsymbol{x}^* = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{y}$$

.

5.2.3　最小二乘问题的求解方法

正则化法

定理 5.2.3. $\boldsymbol{x} \in \mathbb{S}_{LS}$ 的充要条件是 $\boldsymbol{x}$ 为正规方程组(5.21)的解

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b} \tag{5.21}$$

证明. 设 $\boldsymbol{x} \in \mathbb{S}_{LS}$，由定理5.2.2证明知 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}_1$，其中 $\boldsymbol{b}_1 \in \mathbb{R}(\boldsymbol{A})$，而且 $\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} - \boldsymbol{b}_1 = \boldsymbol{b}_2 \in \mathbb{R}(\boldsymbol{A})^{\perp}$. 所以 $\boldsymbol{A}^T\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{A}^T\boldsymbol{b}_2 = 0$. 将 $\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$ 代入 $\boldsymbol{A}^T\boldsymbol{r}(\boldsymbol{x}) = 0$ 即得定理5.21。反之，设 $\boldsymbol{x} \in \mathbb{R}^n$ 满足 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$，则对任意的 $\boldsymbol{y} \in \mathbb{R}^n$ 有

$$\begin{aligned}\|\boldsymbol{b} - \boldsymbol{A}(\boldsymbol{x}+\boldsymbol{y})\|_2^2 &= \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2 - 2\boldsymbol{y}^T\boldsymbol{A}^T(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) + \|\boldsymbol{A}\boldsymbol{y}\|_2^2 \\ &= \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \|\boldsymbol{A}\boldsymbol{y}\|_2^2 \\ &\geqslant \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2\end{aligned}$$

由此即得 $\boldsymbol{x} \in \mathbb{S}_{LS}$。 □

定理 5.2.4. 设 $\boldsymbol{A} \in \mathbb{R}^{m\times n}(m \geq n)$，当 $\boldsymbol{A}$ 的列满秩，$\boldsymbol{A}^T\boldsymbol{A}$ 对称正定可逆，最小二乘问题的解是唯一的，表达式为 $\boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$。其中 $(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$ 又称为 $\boldsymbol{A}$ 的广义逆，用符号 $\boldsymbol{A}^{\dagger}$ 表示。

证明. 因为 $\boldsymbol{A}$ 列满秩，所以 $\boldsymbol{A}^T\boldsymbol{A}$ 可逆，正规方程 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$ 两边同时乘 $(\boldsymbol{A}^T\boldsymbol{A})^{-1}$ 得到 $\boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$。 □

矩阵的广义逆

不是所有的矩阵都有逆矩阵，只有满秩的**方阵**才有逆矩阵，即 $\mathrm{rank}(\boldsymbol{A}) = n$，那么对于不是方阵，不满秩矩阵的情况，数学中也引申出类似于逆矩阵的定义：

定义 5.2.2. 设 $\boldsymbol{A} = (a_{ij})_{m\times n}$ 是数域 $\mathbb{R}$ 上的矩阵, 如果有**唯一**一个矩阵 $\boldsymbol{X}$，满足

$$\begin{cases}\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} = \boldsymbol{A}, \\ \boldsymbol{X}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}, \\ \boldsymbol{A}\boldsymbol{X} = (\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}, \\ \boldsymbol{X}\boldsymbol{A} = (\boldsymbol{X}\boldsymbol{A})^{\mathrm{T}}\end{cases}$$

则矩阵 $\boldsymbol{X}$ 称为矩阵 $\boldsymbol{A}$ 的广义逆, 记作 $\boldsymbol{A}^{+}$.

例 5.2.1. 设矩阵 $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix}1 & 2 & 3 \\ 0 & 1 & 2\end{pmatrix}$$

则矩阵 $\boldsymbol{A}$ 的广义逆为

$$\boldsymbol{A}^{+} = \frac{1}{6}\begin{pmatrix}5 & -8 \\ 2 & -2 \\ -1 & 4\end{pmatrix}$$

广义逆扩充了逆矩阵的范围，将逆矩阵扩大到不是方阵，不满秩的情况。特别地，如果矩阵 $\boldsymbol{A}$ 本身有逆矩阵 $\boldsymbol{A}^{-1}$，则 $\boldsymbol{A}^{-1}$ 显然满足式 (2.1)，即 $\boldsymbol{A}^{+} = \boldsymbol{A}^{-1}$。因此，矩阵 $\boldsymbol{A}$ 的广义逆矩阵 $\boldsymbol{A}^{+}$ 可视为 $\boldsymbol{A}$ 的逆 $\boldsymbol{A}^{-1}$ 的推广。

当 $\boldsymbol{A}$ 列满秩时，我们可以使用 Cholesky 分解法来求解正规方程组。基本步骤如下：

(1) 计算 $\boldsymbol{C}=\boldsymbol{A}^T\boldsymbol{A}$，$\boldsymbol{d}=\boldsymbol{A}^T\boldsymbol{b}$;

(2) 用平方根法计算 $\boldsymbol{C}$ 的 Cholesky 分解：$\boldsymbol{C}=\boldsymbol{L}\boldsymbol{L}^T$;

(3) 求解三角方程组 $\boldsymbol{L}\boldsymbol{y}=\boldsymbol{d}$ 和 $\boldsymbol{L}^T\boldsymbol{x}=\boldsymbol{y}$。

Cholesky 分解法求解的运算量 (乘法次数) 大约为 $mn^2+\frac{1}{3}n^3+O(n^2)$，其中 mn^2 对应计算 $\boldsymbol{A}^T\boldsymbol{A}$，$\frac{1}{3}n^3$ 对应 n 矩阵的 Cholesky 分解，$O(n^2)$ 对应求解三角方程组。

QR 分解法

设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}(m\geq n)$ 列满秩，$\boldsymbol{b}\in\mathbb{R}^m$，将矩阵 $\boldsymbol{A}$ 分解成列正交矩阵 $\boldsymbol{Q}$ 和上三角矩阵 $\boldsymbol{R}$ 的乘积，即

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

然后将列正交矩阵 $\boldsymbol{Q}$ 扩充为正交矩阵 $\left[\boldsymbol{Q},\hat{\boldsymbol{Q}}\right]\in\mathbb{R}^{m\times m}$，得到

$$\boldsymbol{A}=\left[\boldsymbol{Q},\hat{\boldsymbol{Q}}\right]\begin{bmatrix}\boldsymbol{R}\\ \boldsymbol{O}\end{bmatrix}$$

根据 2 范数的正交不变性得

$$\begin{aligned}||\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}||_2^2&=\left\|\left[\boldsymbol{Q},\hat{\boldsymbol{Q}}\right]^T(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})\right\|_2^2\\&=\left\|\left[\boldsymbol{Q},\hat{\boldsymbol{Q}}\right]^T(\boldsymbol{Q}\boldsymbol{R}\boldsymbol{x}-\boldsymbol{b})\right\|_2^2\\&=\left\|\begin{bmatrix}\boldsymbol{R}\boldsymbol{x}-\boldsymbol{Q}^T\boldsymbol{b}\\ -\hat{\boldsymbol{Q}}^T\boldsymbol{b}\end{bmatrix}\right\|_2^2\\&=\left\|\boldsymbol{R}\boldsymbol{x}-\boldsymbol{Q}^T\boldsymbol{b}\right\|_2^2+\left\|\hat{\boldsymbol{Q}}^T\boldsymbol{b}\right\|_2^2\\&\geq\left\|\hat{\boldsymbol{Q}}^T\boldsymbol{b}\right\|_2^2\end{aligned}$$

当且仅当 $\boldsymbol{R}\boldsymbol{x}=\boldsymbol{Q}^T\boldsymbol{b}$ 时等号成立，又因为 $\boldsymbol{R}$ 可逆，所以最小二乘问题的解为

$$\boldsymbol{x}=\boldsymbol{R}^{-1}\boldsymbol{Q}^T\boldsymbol{b}.$$

QR 分解法求解最小二乘问题的基本步骤如下：

(1) 计算 $\boldsymbol{A}$ 的 QR 分解：$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$;

(2) 计算 $\boldsymbol{c}=\boldsymbol{Q}^T\boldsymbol{b}$;

(3) 求解上三角方程组 $\boldsymbol{R}\boldsymbol{x}=\boldsymbol{c}$。

在矩阵分解部分，我们介绍过 Gram-Schmidt 正交化、Hoseholder 变换、Givens 变换三种方法进行 QR 分解。

在计算机中一般使用基于 Householder 变换的 QR 分解，该算法有良好地数值性态，结果通常要比正则化方法精确。但是运算量也比较大，大约为 $2mn^2 - \frac{2}{3}n^3$。

我们也可以使用 Givens 变换来实现 QR 分解，所需的运算量大约是 Householder 方法地两倍，但是如果 $\boldsymbol{A}$ 有较多的零元素，则灵活地使用 Givens 变换会使运算量大为减少。

奇异值分解法

设 $\boldsymbol{A} \in \mathbb{R}^{m\times n}(m \geq n)$ 列满秩，$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{\Sigma} \\ \boldsymbol{O} \end{bmatrix}\boldsymbol{V}^T$ 是 $\boldsymbol{A}$ 的奇异值分解，令 $\boldsymbol{U}_n$ 为 $\boldsymbol{U}$ 的前 n 列组成的矩阵，即 $\boldsymbol{U} = \left[\boldsymbol{U}_n, \tilde{\boldsymbol{U}}\right]$，其中 $\boldsymbol{U}$ 是正交矩阵，根据 2 范数的正交不变性得

$$
\begin{aligned}
\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 &= \left\| \boldsymbol{U}\begin{bmatrix} \boldsymbol{\Sigma} \\ \boldsymbol{O} \end{bmatrix}\boldsymbol{V}^T\boldsymbol{x} - \boldsymbol{b} \right\|_2^2 \\
&= \left\| \begin{bmatrix} \boldsymbol{\Sigma} \\ \boldsymbol{O} \end{bmatrix}\boldsymbol{V}^T\boldsymbol{x} - \left[\boldsymbol{U}_n, \tilde{\boldsymbol{U}}\right]^T\boldsymbol{b} \right\|_2^2 \\
&= \left\| \begin{bmatrix} \boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{x} - \boldsymbol{U}_n^T\boldsymbol{b} \\ -\tilde{\boldsymbol{U}}^T\boldsymbol{b} \end{bmatrix} \right\|_2^2 \\
&= \left\| \boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{x} - \boldsymbol{U}_n^T\boldsymbol{b} \right\|_2^2 + \left\| \tilde{\boldsymbol{U}}^T\boldsymbol{b} \right\|_2^2 \\
&\geq \left\| \tilde{\boldsymbol{U}}^T\boldsymbol{b} \right\|_2^2
\end{aligned}
$$

等号当且仅当 $\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{x} - \boldsymbol{U}_n^T\boldsymbol{b} = 0$ 时成立，即

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

5.2.4 最小二乘问题的变体

对最小二乘问题做一些修改，会得到其他形式的最小二乘问题。

加权最小二乘

在普通的最小二乘法中，我们想要最小化误差向量各项的平方和：

$$
\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 = \sum_{i=1}^{m} r_i^2, \quad r_i = \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} - y_i
$$

其中 $\boldsymbol{a}_i^{\mathrm{T}}, i = 1, \cdots, m$ 是 $\boldsymbol{A}$ 的各列。但是，在某些情形下，方程的残差项并不是同样重要的，相比其他方程，有可能满足某一个方程更重要，这样，我们需要在残差项赋予权重:

$$
f_0(\boldsymbol{x}) = \sum_{i=1}^{m} w_i^2 r_i^2,
$$

其中 $w_i \geq 0$ 是给定的权重。

这样最小化目标函数重写为：

$$f_0(\boldsymbol{x}) = \|\boldsymbol{W}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y})\|_2^2 = \|\boldsymbol{A}_w\boldsymbol{x} - \boldsymbol{y}_w\|_2^2$$

其中

$$\boldsymbol{W} = diag(w_1, \cdots, w_m), \boldsymbol{A}_w \doteq \boldsymbol{W}\boldsymbol{A}, \boldsymbol{y}_w \doteq \boldsymbol{W}\boldsymbol{y}$$

加权最小二乘仍然是普通最小二乘的形式，其权重最小二乘解为：

$$\begin{aligned}\hat{\boldsymbol{x}}_{\text{WLS}} &= (\boldsymbol{A}_w^{\mathrm{T}}\boldsymbol{A}_w)^{-1}\boldsymbol{A}_w^{\mathrm{T}}\boldsymbol{y}_w \\ &= (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y}\end{aligned}$$

约束最小二乘

考虑带有约束的最小二乘问题

$$\begin{aligned}&\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \\ &s.t. \quad \boldsymbol{B}\boldsymbol{x} = \boldsymbol{f}\end{aligned}$$

其中 $\boldsymbol{B}\boldsymbol{x} = \boldsymbol{f}$ 是约束条件。求解需要凸优化知识，在这里只列出解。如果 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 非奇异，且 $\boldsymbol{B}$ 行满秩，则

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

其中 $\boldsymbol{\lambda} = \left[\boldsymbol{B}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{B}^{\mathrm{T}}\right]^{-1}\left[\boldsymbol{B}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{f}\right]$.

总体最小二乘

考虑得到的数据矩阵和数据向量 $\boldsymbol{A}, \boldsymbol{b}$ 都有误差，设实际观测的数据矩阵和数据向量

$$\boldsymbol{A} = \boldsymbol{A}_0 + \boldsymbol{E}, \quad \boldsymbol{b} = \boldsymbol{b}_0 + \boldsymbol{e}$$

其中 $\boldsymbol{E}$ 和 $\boldsymbol{e}$ 分别表示误差数据矩阵和误差数据向量。总体最小二乘的基本思想是：不仅用校正向量 $\Delta\boldsymbol{b}$ 去干扰数据向量 $\boldsymbol{b}$，同时用校正矩阵 $\Delta\boldsymbol{A}$ 去干扰数据矩阵 $\boldsymbol{A}$，以便对 $\boldsymbol{A}$ 和 $\boldsymbol{b}$ 二者内存在的误差或噪声进行联合补偿

$$\boldsymbol{b} + \Delta\boldsymbol{b} = \boldsymbol{b}_0 + \boldsymbol{e} + \Delta\boldsymbol{b} \to \boldsymbol{b}_0$$

$$\boldsymbol{A} + \Delta\boldsymbol{A} = \boldsymbol{A}_0 + \boldsymbol{E} + \Delta\boldsymbol{A} \to \boldsymbol{A}_0$$

以抑制观测误差或噪声对矩阵方程求解的影响，从而实现从有误差的矩阵方程到精确矩阵方程的求解的转换

$$(\boldsymbol{A} + \Delta\boldsymbol{A})\boldsymbol{x} = \boldsymbol{b} + \Delta\boldsymbol{b} \implies \boldsymbol{A}_0\boldsymbol{x} = \boldsymbol{b}_0 \tag{5.22}$$

自然地，我们希望矫正数据矩阵和校正数据向量都尽量小。因此，总体最小二乘问题可以用约束优化问题叙述为：

$$\begin{aligned}\text{TLS:} \quad &\min_{\Delta\boldsymbol{A}, \Delta\boldsymbol{b}, \boldsymbol{x}} \|[\Delta\boldsymbol{A}, \Delta\boldsymbol{b}]\|_{\mathrm{F}}^2 = \|\Delta\boldsymbol{A}\|_{\mathrm{F}}^2 + \|\Delta\boldsymbol{b}\|_2^2 \\ &\text{s.t. } (\boldsymbol{A} + \Delta\boldsymbol{A})\boldsymbol{x} = \boldsymbol{b} + \Delta\boldsymbol{b}\end{aligned}$$

约束条件有时也表示为 $(\boldsymbol{b}+\Delta\boldsymbol{b}) \in \text{Range}\,(\boldsymbol{A}+\Delta\boldsymbol{A})$

由(5.22)，校正过方程的解满足：

$$([\boldsymbol{A},\boldsymbol{b}]+[\Delta\boldsymbol{A},\Delta\boldsymbol{b}])\begin{bmatrix}\boldsymbol{x}\\-1\end{bmatrix}=\boldsymbol{0} \tag{5.23}$$

如果 $([\boldsymbol{A},\boldsymbol{b}]+[\Delta\boldsymbol{A},\Delta\boldsymbol{b}])$ 是列满秩的矩阵，记 $\hat{\boldsymbol{x}}=\begin{bmatrix}\boldsymbol{x}\\-1\end{bmatrix}$，则以 $\hat{\boldsymbol{x}}$ 为未知量的方程：

$$([\boldsymbol{A},\boldsymbol{b}]+[\Delta\boldsymbol{A},\Delta\boldsymbol{b}])\hat{\boldsymbol{x}}=\boldsymbol{0} \tag{5.24}$$

只有零解，与 $\hat{\boldsymbol{x}}$ 的最后一个分量为 -1 矛盾。因此，$([\boldsymbol{A},\boldsymbol{b}]+[\Delta\boldsymbol{A},\Delta\boldsymbol{b}])$ 是一个列亏损矩阵。问题转化为求一个最接近 $[\boldsymbol{A},\boldsymbol{b}]$ 的列亏损矩阵。设 $[\boldsymbol{A},\boldsymbol{b}]$ 的奇异值分解为

$$[\boldsymbol{A},\boldsymbol{b}]=\sum_{i=1}^{n+1}\sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^{\mathrm{T}}$$

其中 σ_i 为 $[\boldsymbol{A},\boldsymbol{b}]$ 的第 i 个奇异值，$\boldsymbol{u}_i,\boldsymbol{v}_i$ 分别为对应的左右奇异向量。

也就是 $[\Delta\boldsymbol{A},\Delta\boldsymbol{b}]=\sigma_{n+1}\boldsymbol{u}_{n+1}\boldsymbol{v}_{n+1}^{\mathrm{T}}$.

设 $\boldsymbol{A}\in\mathbb{R}^{m\times n},\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{b}\in\mathbb{R}^m$，其解为：

$$\hat{\boldsymbol{x}}_{\mathrm{TLS}}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}-\sigma_{n+1}^2\boldsymbol{I})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$$

其中 σ_{n+1} 为 $[-\boldsymbol{b},\boldsymbol{A}]$ 的第 $n+1$ 个奇异值，$\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_{n+1}$。

5.3 特征值计算

在数据科学中，我们一般只讨论实矩阵的特征值问题。

应注意，实矩阵的特征值和特征向量不一定是实数和实向量，但实特征值一定对应于实特征向量（方程 (2.14) 的解），而一般的复特征值对应的特征向量一定不是实向量。此外，由于特征方程为实系数方程，若一个特征值不是实数，则其复共轭也一定是它的特征值。

对于一个实对称矩阵来说，它的 n 个特征值均为实数，并且存在 n 个正交的实特征向量。

5.3.1 圆盘定理

在很多情况下，我们并不需要确切地知道矩阵的每一个特征值的大小，而是要估计出这个矩阵各个特征值大概的范围。

定理 5.3.1. 圆盘定理 设 $\boldsymbol{A}=(a_{kj})\in\mathbb{C}^{n\times n}$，则：

(1) $\boldsymbol{A}$ 的每一个特征值必属于 $\boldsymbol{A}$ 的格什戈林圆盘之中，即对任一特征值 λ 必定存在 $k(1\leq k\leq n)$, 使得

$$|\lambda - a_{kk}| \leq \sum_{j=1,j\neq k}^{n} |a_{kj}| \tag{5.25}$$

用集合的关系来说明，这意味着 $\lambda(\boldsymbol{A}) \subseteq \cup_{k=1}^{n} D_k$，其中 $D_k = \{z||z-a_{kk}| \leq \sum_{j=1,j\neq k}^{n} |a_{kj}|\}$

(2) 若 $\boldsymbol{A}$ 的格什戈林圆盘中有 m 个圆盘组成一连通并集 S, 且 S 与余下的 $n-m$ 个圆盘分离，则 S 内恰好包含 $\boldsymbol{A}$ 的 m 个特征值 (重特征值按重数计)。

下面对定理5.3.1的结论 (1) 进行证明，结论 (2) 的证明超出了本书的范围。

证明. 设 λ 为 $\boldsymbol{A}$ 的任一特征值，则有 $\boldsymbol{Ax} = \lambda\boldsymbol{x}$。$\boldsymbol{x}$ 为非零常量. 设 $\boldsymbol{x}$ 中第 k 个分量最大，即

$$|x_k| = \max_{1\leq j\leq n} |x_j| > 0,$$

考虑线性方程中第 k 个方程

$$\sum_{j=1}^{n} a_{kj}x_j = \lambda x_k,$$

将其中与 x_k 有关的项移到等号左边，其余移到右边，再两边取模得

$$|\lambda - a_{kk}||x_k| = |\sum_{\substack{j=1\\ j\neq k}}^{n} a_{kj}x_j| \leq \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}||x_j| \leq |x_k| \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}| \tag{5.26}$$

最后一个不等式的推导利用了"x 中第 k 个分量最大" 的假设. 将不等式(5.26)除以 $|x_k|$, 即得到式(5.25), 因此证明了定理5.3.1的结论 (1). 上述证明过程还说明, 若某个特征向量的第 k 个分量的模最大, 则相应的特征值必定属于第 k 个圆盘中. □

还可以按照矩阵的每一列元素定义 n 个圆盘, 对于它们定理5.25仍然成立. 下面的定理是圆盘定理的重要推论, 其证明留给感兴趣的读者.

定理 5.3.2. *设 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, 且 $\boldsymbol{A}$ 的对角元均大于 0, 则*

(1) 若 $\boldsymbol{A}$ 严格对角占优, 则 $\boldsymbol{A}$ 的特征值的实部都大于 0.

(2) 若 $\boldsymbol{A}$ 为对角占优的对称矩阵, 则 $\boldsymbol{A}$ 一定是对称半正定矩阵, 若同时 $\boldsymbol{A}$ 非奇异, 则 $\boldsymbol{A}$ 为对称正定矩阵.

例 5.3.1. 试估计矩阵

$$\begin{pmatrix} 4 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -4 \end{pmatrix}$$

的特征值范围。

解. 直接应用圆盘定理, 该矩阵的三个圆盘如下:

$$D_1 : |\lambda - 4| \leq 1, \quad D_2 : |\lambda| \leq 2, \quad D_3 : |\lambda + 4| \leq 2.$$

D_1 与其他圆盘分离, 则它仅含一个特征值, 且必定为实数 *(*若为虚数则其共轭也是特征值，这与 D_1 仅含一个特征值矛盾*)*。所以对矩阵特征值的范围的估计是

$$3 \leq \lambda_1 \leq 5, \quad \lambda_2, \lambda_3 \in D_2 \cup D_3.$$

再对矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 应用圆盘定理，则可以进一步优化上述结果。矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 对应的三个圆盘为

$$D_1^{'} : |\lambda - 4| \leq 2, \quad D_2^{'} : |\lambda| \leq 2, \quad D_3^{'} : |\lambda + 4| \leq 1.$$

这说明 $D_3^{'}$ 中存在一个特征值，且为实数，它属于区间 $[-5, -3]$，经过综合分析可知三个特征值均为实数，它们的范围是

$$\lambda_1 \in [3, 5], \quad \lambda_2 \in [-2, 2], \quad \lambda_3 \in [-5, -3].$$

事实上, 使用 *MATLAB* 的 *eig* 命令可求出矩阵 $\boldsymbol{A}$ 的特征值为 *4.2030, -0.4429, -3.7601.*

在估计特征值范围的时候，我们希望各个圆盘的半径越小越好。所以我们可以通过对矩阵 $\boldsymbol{A}$ 做相似变换, 例如取 $\boldsymbol{X}$ 为对角阵, 然后再应用圆盘定理估计特征值的范围.

例 5.3.2. *(*特征值范围的估计*)*: 选取适当的矩阵 $\boldsymbol{X}$, 应用定理*5.3.1*估计例*5.3.1*中矩阵的特征值范围.

解. 取

$$\boldsymbol{X}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.9 \end{pmatrix}$$

则

$$\boldsymbol{A}_1 = \boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 0 & -\frac{10}{9} \\ 0.9 & 0.9 & -4 \end{pmatrix}$$

的特征值与 $\boldsymbol{A}$ 的相同。对 $\boldsymbol{A}_1$ 应用圆盘定理，得到三个分离的圆盘，它们分别包含一个实特征值，由此得到特征值的范围估计

$$\lambda_1 \in [3, 5], \lambda_2 \in \left[-\frac{19}{9}, \frac{19}{9}\right], \lambda_3 \in [-5.8, -2.2].$$

此外，还可以进一步估计 $\rho(\boldsymbol{A})$ 的范围，即 $3 \leq \rho(\boldsymbol{A}) \leq 5.8$。

上述例子表明，综合运用圆盘定理和矩阵特征值的性质，可对特征值的范围进行一定的估计. 对具体例子，可适当设置相似变换矩阵，尽可能让圆盘相互分离，从而提高估计的有效性。

5.3.2　幂法

幂法是通过求矩阵的特征向量来求出特征值的一种迭代法。它主要用来求按模最大的特征值和相应的特征向量。其优点是算法简单，容易实现，缺点是收敛速度慢，其有效性依赖于矩阵特征值的分布情况。

适于使用幂法的常见情形是：$\boldsymbol{A}$ 的特征值可按模的大小排列为 $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$，且其对应特征向量 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$ 线性无关。此时，任意非零向量 $\boldsymbol{x}^{(0)}$ 均可用 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$ 线性表示，即

$$\boldsymbol{x}^{(0)} = \alpha_1\boldsymbol{\xi}_1 + \alpha_2\boldsymbol{\xi}_2 + \cdots + \alpha_n\boldsymbol{\xi}_n$$

且 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 不全为零。做向量序列 $\boldsymbol{x}^{(k)} = \boldsymbol{A}^k\boldsymbol{x}^{(0)}$, 则

$$\begin{aligned} x^{(k)} &= \boldsymbol{A}^k\boldsymbol{x}^{(0)} = \alpha_1\boldsymbol{A}^k\boldsymbol{\xi}_1 + \alpha_2\boldsymbol{A}^k\boldsymbol{\xi}_2 + \cdots + \alpha_n\boldsymbol{A}^k\boldsymbol{\xi}_n \\ &= \alpha_1\lambda_1^k\boldsymbol{\xi}_1 + \alpha_2\lambda_2^k\boldsymbol{\xi}_2 + \cdots + \alpha_n\lambda_n^k\boldsymbol{\xi}_n \\ &= \lambda_1^k[\alpha_1\boldsymbol{\xi}_1 + \alpha_2(\frac{\lambda_2}{\lambda_1})^k\boldsymbol{\xi}_2 + \cdots + \alpha_n(\frac{\lambda_n}{\lambda_1})^k\boldsymbol{\xi}_n] \end{aligned}$$

由此可见，若 $\alpha_1 \neq 0$，则有

$$\lim_{k=0}^{\inf}(\frac{\lambda_i}{\lambda_1})^k = 0, i = 2, \ldots, n$$

故当 k 充分大的时候，必有

$$\boldsymbol{x}^{(k)} \approx \lambda_1^k\alpha_1\boldsymbol{\xi}_1$$

即 $\boldsymbol{x}^{(k)}$ 可以近似看成 λ_1 对应的特征向量，而 $\boldsymbol{x}^{(k)}$ 与 $\boldsymbol{x}^{(k-1)}$ 分量之比为

$$\frac{\boldsymbol{x}^{(k)}}{\boldsymbol{x}^{(k-1)}} \approx \frac{\lambda_1^k\alpha_1\boldsymbol{\xi}_1}{\lambda_1^{k-1}\alpha_1\boldsymbol{\xi}_1} = \lambda_1$$

于是利用向量序列 $\{\boldsymbol{x}^{(k)}\}$ 即可求出按模最大的特征值 λ_1，又可以求出对应的特征向量 $\boldsymbol{\xi}_1$。

在实际计算中，考虑到当 $|\lambda_1| > 1$ 时，$\lambda_1^k \to \inf$；$|\lambda_1| < 1$ 时，$\lambda_1^k \to 0$，因而计算 $\boldsymbol{x}^{(k)}$ 时可能会发生上溢或者下溢，故每一步将 $\boldsymbol{x}^{(k)}$ 归一化处理，即将 $\boldsymbol{x}^{(k)}$ 的各分量都除以模最大的分量，使 $\left\|\boldsymbol{x}^{(k)}\right\| = 1$，于是求 $\boldsymbol{A}$ 按模最大的特征值 λ_1 和对应的特征向量 ξ_1 的算法，可归纳为如下步骤。

(1) 输入矩阵 $\boldsymbol{A}$，初始向量 $\boldsymbol{v}^{(0)}$，误差限 ϵ，最大迭代次数 N。记 m_0 是 $\boldsymbol{v}^{(0)}$ 按模最大的分量，$\boldsymbol{x}^{(0)} = \boldsymbol{v}^{(0)}/m_0$。置 $k = 0$

(2) 计算 $\boldsymbol{v}^{(k+1)} = \boldsymbol{A}\boldsymbol{x}^{(k)}$。记 m_{k+1} 是 $\boldsymbol{v}^{(k+1)}$ 按模最大的分量，$\boldsymbol{x}^{(k+1)} = \boldsymbol{v}^{(k+1)}/m_{k+1}$

(3) 若 $|m_{k+1} - m_k < \epsilon|$，停止计算，输出近似特征值 m_{k+1} 和近似特征向量 $\boldsymbol{x}^{(k+1)}$ 否则转 (4)

(4) 若 $k < N$ 置 $k = k + 1$ 转 (2) 否则输出计算失败信息，停止计算

定理 5.3.3. 设矩阵 $\boldsymbol{A}$ 的特征值可按模的大小排列为 $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$，且对应特征向量 $\boldsymbol{\xi}_1$,$\boldsymbol{\xi}_2$,...,$\boldsymbol{\xi}_n$ 线性无关. 序列 $\{\boldsymbol{x}^{(k)}\}$ 有算法产生，则有

$$\lim_{k\to\infty}\boldsymbol{x}^{(k)} = \frac{\boldsymbol{\xi}_1}{max\{\boldsymbol{\xi}_1\}} = \boldsymbol{\xi}_1^0, \quad \lim_{k\to\infty} m_k = \lambda_1, \tag{5.27}$$

式中：$\boldsymbol{\xi}_1^0$ 为将 $\boldsymbol{\xi}_1$ 归一化后得到的向量;$max\{\boldsymbol{\xi}_1\}$ 为向量 $\boldsymbol{\xi}_1$ 模最大的分量.

证明. 由算法7.1.6的步 2 和步 3 知

$$\boldsymbol{x}^{(k)}=\frac{\boldsymbol{v}^{(k)}}{m_k}=\frac{\boldsymbol{A}\boldsymbol{x}^{(k-1)}}{m_k}=\frac{\boldsymbol{A}^2\boldsymbol{x}^{k-2}}{m_km_{k-1}}=\cdots=\frac{\boldsymbol{A}^k\boldsymbol{x}^{(0)}}{m_km_{k-1}\cdots m_1}.$$

由于 $\boldsymbol{x}^{(k)}$ 的最大分量为 1，即 $\max\{\boldsymbol{x}^{(k)}\}=1$，故

$$m_km_{k-1}\cdots m_1=max\{\boldsymbol{A}^k\boldsymbol{x}^{(0)}\}$$

从而

$$\boldsymbol{x}^{(k)}=\frac{\boldsymbol{A}^k\boldsymbol{x}^{(0)}}{max\{\boldsymbol{A}^k\boldsymbol{x}^{(0)}\}}=\frac{\lambda_1^k[\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^k\boldsymbol{\xi}_i]}{max\{\lambda_1^k[\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^k\boldsymbol{\xi}_i]\}}$$

$$=\frac{\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^k\boldsymbol{\xi}_i}{max\{\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^k\boldsymbol{\xi}_\mathbf{i}\}}$$

可见

$$\lim_{k\to\infty}\boldsymbol{x}^{(k)}=\frac{\alpha_1\boldsymbol{\xi}_1}{max\{\alpha_1\boldsymbol{\xi}_1\}}=\frac{\boldsymbol{\xi}_1}{max\{\boldsymbol{\xi}_1\}}=\boldsymbol{\xi}_1^0.$$

又

$$\boldsymbol{v}^{(k)}=\boldsymbol{A}\boldsymbol{x}^{(k-1)}=\frac{\boldsymbol{A}^k\boldsymbol{x}^{(0)}}{m_{k-1}\cdots m_1}=\frac{\boldsymbol{A}^k\boldsymbol{x}^{(0)}}{max\{\boldsymbol{A}^{(k-1)}\boldsymbol{x}^{(0)}\}}$$

$$=\frac{\lambda_1^k[\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^k\boldsymbol{\xi}_i]}{\lambda_1^{k-1}max\{[\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^{k-1}\boldsymbol{\xi}_i]}$$

注意到 m_k 是 $\boldsymbol{v}^{(k)}$ 模的最大的分量，既有

$$m_k=max\{\boldsymbol{v}^{(k)}\}=\lambda_1\frac{max\{\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^k\boldsymbol{\xi}_i\}}{max\{\alpha_1\boldsymbol{\xi}_1+\sum_{i=2}^{n}\alpha_i(\frac{\lambda_i}{\lambda_1})^{k-1}\boldsymbol{\xi}_i\}}$$

从而 $\lim_{k\to\infty}m_k=\lambda_1$ 成立. 证毕. □

幂法的收敛速度与比值 $|\lambda_2/\lambda_1|$ 的大小有关，$|\lambda_2/\lambda_1|$ 越小，收敛速度越快，当此比值接近于 1 时，收敛速度是非常缓慢的。因此可以对原矩阵作原点位移，令

$$\boldsymbol{B}=\boldsymbol{A}-\alpha\boldsymbol{I}$$

式中 α 为参数。选择此参数可使矩阵 $\boldsymbol{B}$ 的上述比值更小，以加快幂法的收敛速度。设矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$，对应的特征向量为 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$，则矩阵 $\boldsymbol{B}$ 的特征值为 $\lambda_1 - \alpha, \lambda_2 - \alpha, \ldots, \lambda_n - \alpha$，$\boldsymbol{B}$ 的特征向量与 $\boldsymbol{A}$ 的特征向量相同。假设原点位移后，$\boldsymbol{B}$ 的特征值 $\lambda_1 - \alpha$ 仍为模最大的特征值，选择 α 的目的是使

$$\max_{2 \le i \le n} \frac{\lambda_i - \alpha}{\lambda_1 - \alpha} < \frac{\lambda_2}{\lambda_1}$$

适当地选择 α 可使幂法的收敛速度得到加速。此时 $m_k \to \lambda_1 - \alpha, m_k + \alpha \to \lambda_1$，而 $\boldsymbol{x}^{(K)}$ 仍然收敛于 $\boldsymbol{A}$ 的特征向量 $\boldsymbol{\xi}_1^0$。这种加速方法叫做原点位移法。

在实际计算中，由于矩阵的特征值分布情况事先一般是不知道的，参数 α 的选取存在困难，因为 α 的选取要保证 $\lambda_1 - \alpha$ 仍然是矩阵 $\boldsymbol{B}(= A - \alpha \boldsymbol{I})$ 模最大的特征值，故原点位移法是很难实现的。但是在反幂法中，原点位移参数 α 是非常容易选取的，因此，带原点位移的反幂法已经成为改进特征值和特征向量精度的标准算法。

5.3.3　反幂法

设 $\boldsymbol{A}$ 可逆，则对 $\boldsymbol{A}$ 的逆阵 $\boldsymbol{A}^{-1}$ 施以幂法称为反幂法。由于 $\boldsymbol{A}\boldsymbol{\xi}_i = \lambda_i \boldsymbol{\xi}_i$ 时，成立 $\boldsymbol{A}^{-1}\boldsymbol{\xi}_i = \lambda_i^{-1}\boldsymbol{\xi}_i$。因此，若 $|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_{n-1}| > |\lambda_n|$，则 λ_n^{-1} 是 $\boldsymbol{A}^{-1}$ 按模最大的特征值，此时按反幂法，必有

$$m_k \to \lambda_n^{-1}, \boldsymbol{x}^{(k)} \to \boldsymbol{\xi}_n^0$$

其收敛率为 $|\lambda_n / \lambda_{n-1}|$。任取初始向量 $\boldsymbol{x}^{(0)}$，构造向量序列

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{A}^{-1}\boldsymbol{x}^{(k)}, k = 0, 1, 2, \ldots$$

按幂法计算即可，但用上述式子计算，首先要求 $\boldsymbol{A}^{-1}$，这比较麻烦而且效率不高，实际计算中通常用解方程组的办法，即用

$$\boldsymbol{A}\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}, k = 0, 1, 2, \ldots$$

求 $\boldsymbol{x}^{(k+1)}$。为防止计算机溢出，实际计算时所用公式为

$$\begin{aligned} \boldsymbol{v}^{(k)} &= \boldsymbol{x}^{(k)} / \max(\boldsymbol{x}^{(k)}), \\ \boldsymbol{A}\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)}, \end{aligned} \quad k = 0, 1, 2, \ldots$$

式中：$\max(\boldsymbol{x}^{(k)})$ 为 $\boldsymbol{x}^{(k)}$ 模最大的分量。

若 $\boldsymbol{A}$ 的特征值是 λ，则 $\lambda - \alpha$ 是 $\boldsymbol{A} - \alpha \boldsymbol{I}$ 的特征值。因此反幂法可以用于已知矩阵的近似特征值为 α 时，求矩阵的特征向量并且提高特征值精度。

此时，可以用原点位移法来加速迭代过程，于是上式相应为

$$\boldsymbol{A}\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)}, k = 0, 1, 2, \ldots$$

求 $\boldsymbol{x}^{(k+1)}$。为防止计算机溢出，实际计算时所用公式为

$$\begin{aligned} \boldsymbol{v}^{(k)} &= \boldsymbol{x}^{(k)} / \max(\boldsymbol{x}^{(k)}), \\ (\boldsymbol{A} - \alpha \boldsymbol{I})\boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)}, \end{aligned} \quad k = 0, 1, 2, \ldots$$

算法（反幂法）

(1) 选取初值 $\boldsymbol{x}^{(0)}$，近似值 α，误差限 ϵ，最大迭代次数 N。记 m_0 为 $\boldsymbol{x}(0)$ 中按模最大的分量，$\boldsymbol{v}^{(0)}=\boldsymbol{x}^{(0)}/m_0$。置 $k=0$

(2) 解方程组 $(\boldsymbol{A}-\alpha\boldsymbol{I})\boldsymbol{x}^{(k+1)}=\boldsymbol{v}^{(k+1)}$。

(3) 记 m_{k+1} 为 $\boldsymbol{x}^{(k+1)}$ 中按模最大的分量，$\boldsymbol{v}^{(k+1)}=\boldsymbol{x}^{(k+1)}/m_{k+1}$。

(4) 若 $|m_{k+1}^{-1}-m_k^{-1}|<\epsilon$，则置 $\lambda=m_{k+1}^{-1}+\alpha$，输出 λ 和 $\boldsymbol{x}^{(k+1)}$，停止计算；否则，转 (5)

(5) 若 $k<N$，置 $k=k+1$，转 (2)，否则输出计算失败信息，停止计算。

5.3.4 特征值计算的应用：Pagerank 网页排名

接下来我们将介绍一个用于网页排名的算法——Pagerank。

(1) **问题背景** 互联网（Internet）的使用已经深入到人们的日常生活中，其巨大的信息量和强大的功能给生产、生活带来了很大的便利. 随着网络的信息量越来越庞大，如何有效地搜索出用户真正需要的信息变得十分重要。自 1998 年搜索引擎网站 Google 创立以来，网络搜索引擎成为解决上述问题的重要手段。

1998 年，美国斯坦福大学的博士生 Larry Page 和 Sergey Brin 创立了 Google 公司，他们的核心技术就是通过 Pagerank 技术对海量的网页进行重要性分析。该技术利用网页相互链接的关系对网页进行组织，确定出每个网页的重要级别 (Pagerank). 当用户进行搜索时，Google 找出符合搜索要求的网页，并按它们的 Pagerank 大小依次列出。这样，用户一般显示结果的第一页或者前几页就能找到真正有用的结果。

形象地解释，Pagerank 技术的基本原理是：如果网页 A 链接到网页 B, 则认为“网页 A 投了网页 B”一票，而且如果网页 A 是级别高的网页，则网页 B 的级别也相应地高。

(2) **数学问题建模** 假设 n 是 Internet 中所有可访问网页的数目，此数值非常大，再 2010 年已接近 100 亿。定义 $n\times n$ 的网页连接矩阵 $\boldsymbol{G}=(g_{ij})\in\mathbb{R}^{n\times n}$，若从网页 j 有一个链接到网页 i，则 $g_{ij}=1$，否则 $g_{ij}=0$。矩阵 $\boldsymbol{G}$ 有如下特点：

(1) $\boldsymbol{G}$ 矩阵是大规模稀疏矩阵；

(2) 第 j 列非零元素的位置表示了从网页 j 链接出去的所有网页;

(3) 第 i 列非零元素的位置表示了所有链接到网页 i 的网页；

(4) $\boldsymbol{G}$ 中非零元的数目为整个 Internet 中存在的超链接的数量；

(5) 记 $\boldsymbol{G}$ 矩阵行元素之和 $r_i=\sum_j g_{ij}$，它表示第 i 个网页的“入度”；

(6) 记 $\boldsymbol{G}$ 矩阵列元素之和 $c_j=\sum_i g_{ij}$，它表示第 j 个网页的“出度”。

要计算 PageRank，可假设一个随机上网“冲浪”的过程，即每次看完当前网页后，有两种选择：

(1) 在当前网页中随机选一个超链接进入下一个网页；

(2) 随机地新开一个网页；

设 p 为选择当前网页上链接的概率（比如 $p=0.85$），则 $1-p$ 为不选当前网页的链接而随机打开一个网页的概率。若当前网页是网页 j，则如何计算下一步浏览到达网页 i 的概率（网页 j 到 i 的转移概率）？它有两种可能性：

(1) 若网页 i 在网页 j 的链接上，其概率为 $p\cdot 1/c_j+(1-p)\cdot 1/n$；

(2) 若网页 i 不在网页 j 的链接上，其概率为 $(1-p)\cdot 1/n$。

由于网页 i 是否在网页 j 的链接上由 g_{ij} 决定，网页 j 到 i 的转移概率为

$$a_{ij}=g_{ij}\left[p\cdot\frac{1}{c_j}+(1-p)\cdot\frac{1}{n}\right]+(1-g_{ij})\left[(1-p)\cdot\frac{1}{n}\right]=\frac{pg_{ij}}{c_j}+\frac{1-p}{n}$$

应注意的是，若 $c_j=0$ 意味着 $g_{ij}=0$，上式改为 $a_{ij}=1/n$。任意两个网页之间的转移概率形成了一个转移矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$。设矩阵 $\boldsymbol{D}$ 为各个网页出度的倒数（若没有出度，设为 1）构成的 n 阶对角阵，$\boldsymbol{e}$ 为全是 1 的 n 维向量，则

$$\boldsymbol{A}=p\boldsymbol{GD}+\frac{1-p}{n}\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}.$$

这在数学上称为马尔可夫过程。若这样的随机“冲浪”一直进行下去，某个网页被访问的极限概率就是它的 PageRank。

设 $x_i^{(k)},i=1,2,\cdots,n$ 表示某时刻 k 浏览网页 i 的概率 $\left(\sum x_i^{(k)=1}\right)$，向量 $\boldsymbol{x}^{(k)}$ 表示当前时刻浏览个网页的概率分布。那么下一时刻浏览到网页 i 的概率为 $\sum_{j=1}^{n}a_{ij}x_i^{(k)}$，此时浏览个网页的概率分布为 $\boldsymbol{x}^{(k+1)}=\boldsymbol{A}\boldsymbol{x}^{(k)}$。

当这个过程无限进行下去，达到极限情况，即网页访问概率 $\boldsymbol{x}^{(k)}$ 收敛到一个极限值，这个极限向量 $\boldsymbol{x}$ 为个网页的 PageRank，他满足 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}$，且 $\sum_{i=1}^{n}x_i=1$。

总结一下，我们要求解的问题是在给定 $n\times n$ 的网页连接矩阵 $\boldsymbol{G}$, 以及选择当前网页链接的概率 p 时，计算特征值 1 对应的特征向量 $\boldsymbol{x}$

$$\begin{cases}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}\\ \displaystyle\sum_{i=1}^{n}x_i=1\end{cases}$$

易知 $\|\boldsymbol{A}\|_1=1$，所以 $\rho(\boldsymbol{A})\leq 1$。又考虑矩阵 $\boldsymbol{L}=\boldsymbol{I}-\boldsymbol{A}$，容易验证它各列元素和均为 0，则 $\boldsymbol{L}$ 为奇异矩阵，所以 $\det(\boldsymbol{I}-\boldsymbol{A})=0$，1 是 $\boldsymbol{A}$ 的特征值且为主特征值。更进一步，用圆盘定理考察矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的特征值分布，图 (a) 显示了第 j 个圆盘 $D_j(j=1,2,\cdots,n)$，显然其圆心 $a_{jj}>0$，半径 r_j 满足 $a_{jj}+r_j=1$，因此除了 1 这一点，圆盘上任何一点到圆心的距离（即复数的模）都小于 1。这就说明，1 是矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 和 $\boldsymbol{A}$ 的唯一主特征值。对于实际的大规模稀疏矩阵 $\boldsymbol{A}$，幂法是求其主特征向量的可靠的、唯一的选择。

网页的 PageRank 完全由所有网页的超链接结构所决定，隔一段时间重新算一次 PageRank 以反映互联网的发展变化，此时将上一次计算的结果作为幂法的迭代初值可提高收敛速度。由于迭代向量以及矩阵 $\boldsymbol{A}$ 的物理意义。在使用幂法时并不需要对向量进行规格化，而且不需要形成矩阵 $\boldsymbol{A}$。通过遍历整个网页的数据库，根据网页间超链接关系即可得到 $\boldsymbol{A}\boldsymbol{x}^{(k)}$ 的结果。

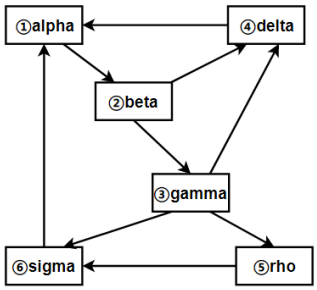

图 5.3: 网页链接关系

用一个只有 6 个网页的微型网络作为例子，其网页链接关系如图5.3所示。通过下述 MATLAB 命令可生成矩阵 ***G***

$i = [2\ 3\ 4\ 4\ 5\ 6\ 1\ 6\ 1];$

$j = [1\ 2\ 2\ 3\ 3\ 3\ 4\ 5\ 6];$

$n = 6;$

$G = aparse(i, j, 1, n, n);$

再使用下述命令得到矩阵 ***A***

$c = full(sum(G));$

$D = spdiags(1./c', 0, n, n);$

$e = ones(n, 1);$

$p = .85; delta = (1 - p)/n;$

$A = p * G * D + delta * e * e';$

得到的矩阵 ***A*** 为

$$\begin{pmatrix} 0.025 & 0.025 & 0.025 & 0.875 & 0.025 & 0.875 \\ 0.875 & 0.025 & 0.025 & 0.025 & 0.025 & 0.025 \\ 0.025 & 0.45 & 0.025 & 0.025 & 0.025 & 0.025 \\ 0.025 & 0.45 & 0.3083 & 0.025 & 0.025 & 0.025 \\ 0.025 & 0.025 & 0.3083 & 0.025 & 0.025 & 0.025 \\ 0.025 & 0.025 & 0.3083 & 0.025 & 0.875 & 0.025 \end{pmatrix}$$

使用幂方法可求出其主特征向量，即 PageRank 为

$$\boldsymbol{x} = [0.2675\ 0.2524\ 0.1323\ 0.1697\ 0.0625\ 0.1156]^{\mathrm{T}}$$

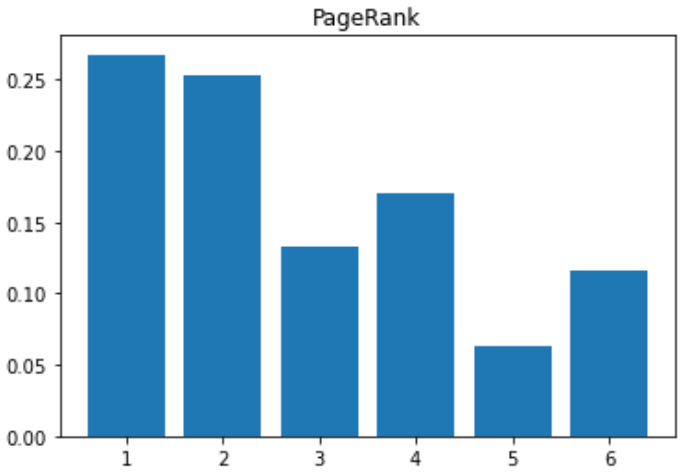

图 5.4: 网页的级别高低

使用 MATLAB 的 bar 命令，将 $\boldsymbol{x}$ 的各分量显示如图5.4所示，从中看出各个网页的级别高低，虽然链接数目一样，但是网页 alpha 1 的链接比 delta 4 和 sigma 6 都高，而由于高级别的 alpha 1 链接在 beta2 上面，受到 alpha 1 的影响，beta 2 的级别第二高。

5.4　阅读材料

本章介绍了数值线性代数三大核心主题内容，包括线性方程的求解、最小二乘问题和特征值的求解。数据科学中的很多问题最终都归结为线性方程的求解，因此这一章主要介绍线性方程组的类型和解的结构，引入线性方程组和最小二乘问题的求解方法，并讨论解的敏感性，这些内容将与后续优化问题求解、数据科学中的线性回归问题相联系。此外，还介绍了大规模矩阵求解特征值的一些计算方法，包括幂迭代法，这已被广泛应用于数据科学中的搜索技术 pagerank 的矩阵特征值计算。此外，用于对高维数据进行非线性降维和聚类的更现代的谱方法，如 Isomap（Tenenbaum 等，2000），Laplacian 特征映射（Belkin 和 Niyogi，2003），Hessian 特征映射（Donoho 和 Grimes，2003），谱聚类（Shi 和 Malik，2000）等，每一个都需要计算正定核的特征向量和特征值，这些核心计算通常由低秩矩阵近似技术（Belabbas 和 Wolfe，2009）支持，正如我们在 SVD 中遇到的那样。另外，关于稠密数值线性代数可参考（Golub 和 Van Loan,1989），（Trefethen 和 Bau, 1997）等。（Gill，Murray,1981）和（Wright,1997），（Wright 以及 Nocedal,1999）等书籍注重于数值优化问题的数值线性代数介绍。关于数值线性代数软件包，可以参考 LAPACK，其包括常规的稠密的线性代数算法的高质量实现。LAPACK 在基本线性代数子程序（BLAS）的基础上建成，后者是基本的向量和矩阵运算的程序库，可以很容易地根据具体的计算机结构的优点进行定制，也可得到求解稀疏的线性方程组的一些源代码，包括 SPOOLES，SuperLU，UMFPACK

以及 WSMP 等等，这里提到的只是其中少数几个。

习题

习题 5.1. 设 $\boldsymbol{A}=\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 用正则化方法求对应的 LS 问题的解。

习题 5.2. 设 $\boldsymbol{A}=\begin{bmatrix}1&3&1&1\\2&0&0&0\\1&0&0&0\end{bmatrix}$，$\boldsymbol{b}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 求对应的 LS 问题的全部解。

习题 5.3. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 且存在 $\boldsymbol{X}\in\mathbb{R}^{n\times m}$ 使得对每一个 $\boldsymbol{b}\in\mathbb{R}^m$,$\boldsymbol{x}=\boldsymbol{X}\boldsymbol{b}$ 均极小化 $\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2$. 证明 $\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}=\boldsymbol{A}$ 和 $(\boldsymbol{A}\boldsymbol{X})^T=\boldsymbol{A}\boldsymbol{X}$.

习题 5.4. 利用等式

$$\|\boldsymbol{A}(\boldsymbol{x}+\alpha\boldsymbol{w})-\boldsymbol{b}\|_2^2=\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2+2\alpha\boldsymbol{w}^T\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})+\alpha^2\|\boldsymbol{A}\boldsymbol{w}\|_2^2$$

证明：如果 $\boldsymbol{x}\in\boldsymbol{X}_{LS}$, 那么 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^T\boldsymbol{b}$

习题 5.5. 给定点集 $\boldsymbol{p}_1,\cdots,\boldsymbol{p}_m\in\mathbb{R}^n$ 构成的 $m\times n$ 矩阵 $\boldsymbol{P}=[\boldsymbol{p}_1,\cdots,\boldsymbol{p}_m]$. 考虑问题

$$\min_{\boldsymbol{X}}F(\boldsymbol{X})=\sum_{i=1}^{m}\|\boldsymbol{x}_i-\boldsymbol{p}_i\|_2^2+\frac{\lambda}{2}\sum_{1\le i,j\le m}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|_2^2$$

其中 $\lambda\ge 0$ 为参数，变量是一个 $m\times n$ 矩阵 $\boldsymbol{X}=[\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m]$，其中 $\boldsymbol{x}_i\in\mathbb{R}^n$ 是 $\boldsymbol{X}$ 的第 i 列，$i=1□\cdots,m$. 上述问题尝试聚类点集 $\boldsymbol{p}_i$，第一项鼓励聚类中心 $\boldsymbol{x}_i$ 靠近对应的点 $\boldsymbol{p}_i$，第二项鼓励 $\boldsymbol{x}_i$ 们之间彼此靠近，当 λ 增大的时候，对应更高的组群影响。

1. 请说明这个问题属于最小二乘类问题。不需要明确阐述这个问题的形式。

2. 证明 $\frac{1}{2}\sum_{1\le i,j\le m}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|_2^2=tr\boldsymbol{X}\boldsymbol{H}\boldsymbol{X}^T$, 其中 $\boldsymbol{H}=m\boldsymbol{I}_m-\boldsymbol{1}\boldsymbol{1}^T$ 是一个 $m\times m$ 矩阵，$\boldsymbol{I}_m$ 是 $m\times m$ 单位矩阵，$\boldsymbol{1}$ 是 $\mathbb{R}^n$ 中的单位向量。

3. 证明 $\boldsymbol{H}$ 是半正定的。

4. 证明函数 F 在矩阵 $\boldsymbol{X}$ 处的梯度是一个 $n\times m$ 矩阵，为：

$$\nabla F(\boldsymbol{X})=2(\boldsymbol{X}-\boldsymbol{P}+\lambda\boldsymbol{X}\boldsymbol{H})$$

提示：对于第二项，找到函数的一阶展式，$\triangle\to trace((\boldsymbol{X}+\triangle)\boldsymbol{H}(\boldsymbol{X}+\triangle)^T)$, 其中 $\triangle\mathbb{R}^{n,m}$。

5. 依据最小二乘问题的最优条件为目标函数的梯度为零。证明最优点集的形式为:

$$\boldsymbol{x}_i=\frac{1}{m\lambda+1}\boldsymbol{p}_i+\frac{m\lambda}{m\lambda+1}\hat{\boldsymbol{p}},i=1,\cdots,m,$$

其中 $\hat{\boldsymbol{p}}=□(1/m)(\boldsymbol{p}_1+\cdots+\boldsymbol{p}_m)$ 是给定点集的中心。

6. 阐述你的结果，你认为这是聚类点集的一个好的模型么？

习题 5.6. 判断 $[1,3,4]$ 的转置是否在 $\boldsymbol{A}$ 的零空间中

$$\boldsymbol{A}=\begin{bmatrix}3 & 5 & -3\\ 6 & -2 & 0\\ -8 & 4 & 1\end{bmatrix}$$

习题 5.7. 求矩阵

$$\begin{bmatrix}5 & 21 & 19\\ 13 & 23 & 2\\ 8 & 14 & 1\end{bmatrix}$$

的行空间和列空间

习题 5.8. 简答：阐述非负矩阵分解和主成分分析的相同点和不同点

参考文献

[1] E.Anderson, Z.Bai, C.Bischof, S.Blackford, J.Demrnel, J.Dongarra, J.DuCroz, A.Greenbaum, S.Hammarling，A.McKenney，and D.Sorensen. LAPACK Users' Guide. Society for Industrial and Applied Mathematics, third edition, 1999. Available from www.netlib.org/lapack.

[2] C.Ashcraft，D.Pierce，D.K.Wah，and J.Wu.The Reference Manual for SPOOLES Version 2.2: An Object Oriented Software Library for Solving Sparse Linear Systems of Equations, 1999. Availablc from www.netlib.org/linalg/spooles/spooles.2.2.html.

[3] T.A.Davis.UMFPACK User Guide, 2003.Available from wv.cise.ufl.edu/research/sparse/umfpack.

[4] J.W.Dexnmel.Applied Numerical Linear Algebra.Society for Industrial and Applied Mathematics, 1997.

[5] I.S.Duff, A.M.Erismaim, and J.K.Reid.Direct Methods for Sparse Matrices.Clarendon Press, 1986.

[6] J.W.Demxnel, J.R.Gilbert, and X.S.Li.SuperLU Users' Guide, 2003.Available from crd.lbl.gov/ xiaoye/SuperLU.

[7] I.S.Duff.The solution of augmented systems.In D.F.Griffiths and G.A.Watson, editors, Numerical Analysis 1993.Proceedings of the 15th Dundee Conference, pages 40- 55.Longman Scientific & Technical, 1993.

[8] G.Golub and C.F.Van Loan.Matrix Computations.Johns Hopkins University Press, second edition, 1989.

[9] A.George and J.W.-H.Liu.Computer sohstion of large sparse positive definite systems.Prentice-Hall, 1981.

[10] Belkin, Mikhail, and Niyogi, Partha. 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural computation, 15(6), 1373–1396.

[11] Tenenbaum, Joshua B, De Silva, Vin, and Langford, John C. 2000. A global geometric framework for nonlinear dimensionality reduction. science, 290(5500), 2319–2323.

[12] Donoho, David L, and Grimes, Carrie. 2003. Hessian eigenmaps: Locally linear em- bedding techniques for high-dimensional data. Proceedings of the National Academy of Sciences, 100(10), 5591–5596.

[13] Belabbas, Mohamed-Ali, and Wolfe, Patrick J. 2009. Spectral methods in machine learning and new strategies for very large datasets. Proceedings of the National Academy of Sciences, pnas–0810600105.

[14] Shi, Jianbo, and Malik, Jitendra. 2000. Normalized cuts and image segmentation. IEEE Transactions on pattern analysis and machine intelligence, 22(8), 888–905.

[15] A.Gupta.WSMP: Watson Sparse Matrix Package.Part I —Direct Solution of Symmetric Sparse Systems.Part II —Direct Solution of General Sparse Systems, 2000.Available from www.cs.umn.edu/ agupta/wsmp.

[16] N.J.Higham.Accuracy and Stability of Numerical Algorithms.Society for Industrial and Applied Mathematics, 1996,

[17] P.E.Gill, W.Murray, and M.H.Wright.Practical Optimization, Acatlemic Press, 1981.

[18] J.Nocedal and S.J.Wright.Numerical OptimizatiorL Springer, 1999.

[19] L.N.Trefethen and D.Bau, III.Numerical Linear Algebra.Society for Industrial and Applied Mathematics, 1997.

[20] S.J.Wright.Primal-Dual Interior-Point Methods.Society for Industrial and Applied Mathematics, 1997.