

因而条件数变为 δ^{-2} 数量级。

另外一方面, 如果我们利用 A 的 QR 分解来解超定方程 $Ax = b$ 的话, 那么由于 $Q^H Q = I$, 所以

$$\text{cond}(Q) = 1 \quad (6.1.11)$$

$$\text{cond}(A) = \text{cond}(AQ) = \text{cond}(A) \quad (6.1.12)$$

此时, b 和 A 的误差的影响将分别如式 (6.1.5) 和式 (6.1.7) 所示, 与 A 的条件数成正比。

以上事实告诉我们, 求解超定方程问题的 QR 分解方法具有比最小二乘方法更好的数值稳定性 (更小的条件数)。

若条件数“很大”, 线性方程问题便称为 (相对于范数 $\|\cdot\|$) 病态的。此时, 对于一接近真实 b 的 b^* , 由于条件数很大, 所以与 b^* 对应的解就会远离对应于 b 的解。解决这类病态问题的一种比 QR 分解更加有效的方法是总体最小二乘法 (将在第 7 章介绍), 它的基础就是下一节要讨论的矩阵的奇异值分解。事实上, 正如以后几节将看到的那样, 矩阵的奇异值分解已被广泛应用于解决工程学科中的许多重要问题。

6.2 奇异值分解

奇异值分解 (singular value decomposition, SVD) 是现代数值分析 (尤其是数值计算) 的最基本和最重要的工具之一。本节介绍奇异值分解的定义、几何解释以及奇异值的性质。

6.2.1 奇异值分解及其解释

奇异值分解最早是由 Beltrami 在 1873 年对实正方矩阵提出来的^[38]。Beltrami 从双线性函数

$$f(x, y) = x^T A y, \quad A \in R^{n \times n}$$

出发, 通过引入线性变换

$$x = U\xi, \quad y = V\eta$$

将双线性函数变为

$$f(x, y) = \xi^T S \eta$$

式中

$$S = U^T A V \quad (6.2.1)$$

Beltrami 观测到, 如果约束 U 和 V 为正交矩阵, 则它们的选择各存在 $n^2 - n$ 个自由度。他提出利用这些自由度使矩阵 S 的对角线以外的元素全部为零, 即矩阵 $S = \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ 为对角矩阵。于是, 用 U 和 V^T 分别左乘和右乘式 (6.2.1), 并利用 U 和 V 的正交性, 立即得到

$$A = U \Sigma V^T \quad (6.2.2)$$

这就是 Beltrami 于 1873 年得到的实正方矩阵的奇异值分解^[38]。1874 年, Jordan 也独立地推导出了实正方矩阵的奇异值分解^[241]。有关奇异值分解的这段发明历史, 可参见 MacDuffee 的书^[294,p.78]或 Stewart 的评述论文^[433]。文献^[433]还详细地评述了奇异值分解的整个早期历史。

后来, Autonne^[22]于 1902 年把奇异值分解推广到复正方矩阵; Eckart 与 Young^[141]于 1939 年又进一步把它推广到一般的长方形矩阵。因此, 现在常将任意复长方矩阵的奇异值分解定理称为 Autonne-Eckart-Young 定理, 详见下述。

定理 6.2.1 (矩阵的奇异值分解) 令 $A \in R^{m \times n}$ (或 $C^{m \times n}$), 则存在正交 (或酉) 矩阵 $U \in R^{m \times m}$ (或 $C^{m \times m}$) 和 $V \in R^{n \times n}$ (或 $C^{n \times n}$) 使得

$$A = U \Sigma V^T \text{ (或 } U \Sigma V^H \text{)} \quad (6.2.3)$$

式中

$$\Sigma = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix} \quad (6.2.4)$$

且 $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, 其对角元素按照顺序

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0, \quad r = \text{rank}(A) \quad (6.2.5)$$

排列。

以上定理最早是 Eckart 与 Young^[141]于 1939 年证明的, 但证明较繁杂。下面采用的是 Klema 与 Laub^[256]的比较简单的证明。

证明 因为 $A^T A \geq 0$, 所以 $\sigma(A^T A) \subseteq [0, +\infty)$ 。记 $\sigma(A^T A) = \{\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2\}$, 并将它们的顺序安排成 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 = \sigma_{r+1} = \dots = \sigma_n$ 。令 v_1, v_2, \dots, v_n 是对应的正交特征向量组, 而且

$$V_1 = [v_1, v_2, \dots, v_r] \quad (6.2.6)$$

$$V_2 = [v_{r+1}, v_{r+2}, \dots, v_n] \quad (6.2.7)$$

于是, 若 $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, 则有 $A^T A V_1 = V_1 \Sigma_1^2$, 由此得到

$$\Sigma_1^{-1} V_1^T A^T A V_1 \Sigma_1^{-1} = I \quad (6.2.8)$$

另有 $A^T A V_2 = V_2 \times O$ 使得 $V_2^T A^T A V_2 = O$, 因此 $A V_2 = O$ 。令 $U_1 = A V_1 \Sigma_1^{-1}$, 则由式 (6.2.8), 我们有 $U_1^T U_1 = I$ 。选择任意 U_2 , 使得 $U = [U_1, U_2]$ 正交。于是

$$U^T A V = \begin{bmatrix} U_1^T A V_1 & U_1^T A V_2 \\ U_2^T A V_1 & U_2^T A V_2 \end{bmatrix} = \begin{bmatrix} \Sigma_1 & O \\ U_2^T U_1 \Sigma_1 & O \end{bmatrix} = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix} = \Sigma$$

从而得到所希望的结果 $A = U \Sigma V^T$ 。 ■

数值 $\sigma_1, \sigma_2, \dots, \sigma_r$ 连同 $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$ 一起称作矩阵 A 的奇异值。

定义 6.2.1 矩阵 $A_{m \times n}$ 的奇异值 σ_i 称为单奇异值, 若 $\sigma_i \neq \sigma_j, \forall j \neq i$ 。

下面是关于奇异值和奇异值分解的几点解释和标记。

(1) $n \times n$ 矩阵 V 为酉矩阵, 用 V 右乘式 (6.2.3), 得 $AV = U\Sigma$, 其列向量形式为

$$Av_i = \begin{cases} \sigma_i u_i, & i = 1, 2, \dots, r \\ 0, & i = r+1, r+2, \dots, n \end{cases} \quad (6.2.9)$$

因此, V 的列向量 v_i 称为矩阵 A 的右奇异向量 (right singular vector), V 称为 A 的右奇异向量矩阵 (right singular vector matrix)。

(2) $m \times m$ 矩阵 U 是酉矩阵, 用 U^H 左乘式 (6.2.3), 得 $U^H A = \Sigma V$, 其列向量形式为

$$u_i^H A = \begin{cases} \sigma_i v_i^T, & i = 1, 2, \dots, r \\ 0, & i = r+1, r+2, \dots, n \end{cases} \quad (6.2.10)$$

因此, U 的列向量 u_i 称为矩阵 A 的左奇异向量 (left singular vector), 并称 U 为 A 的左奇异向量矩阵 (left singular vector matrix)。

(3) 用 u_i^H 左乘式 (6.2.9), 并注意到 $u_i^H u_i = 1$, 易得

$$u_i^H A v_i = \sigma_i, \quad i = 1, 2, \dots, \min\{m, n\} \quad (6.2.11)$$

或用矩阵形式写成

$$U^H A V = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \sigma_n \end{bmatrix} \quad (6.2.12)$$

式 (6.2.3) 和式 (6.2.11) 是矩阵奇异值分解的两种定义方式。事实上, 式 (6.2.3) 很容易由式 (6.2.11) 导出。由于 U 和 V 分别是 $m \times m$ 和 $n \times n$ 酉矩阵, 满足 $UU^H = I_m$ 和 $VV^H = I_n$, 所以在式 (6.2.11) 两边左乘 U 和右乘 V^H 后, 立即得式 (6.2.3)。这也可以看作是定理 6.2.1 的另一种推导。

(4) 矩阵 A 的奇异值分解式 (6.2.3) 可以改写成向量表达形式:

$$A = \sum_{i=1}^r \sigma_i u_i v_i^H \quad (6.2.13)$$

这种表达有时称为 A 的并向量 (奇异值) 分解 (dyadic decomposition) [184]。

(5) 由式 (6.2.3) 易得

$$AA^H = U \Sigma^2 U^H \quad (6.2.14)$$

这表明, $m \times n$ 矩阵 A 的奇异值 σ_i 是矩阵乘积 AA^H 的特征值 (这些特征值是非负的) 的正平方根。

(6) 当矩阵 A 的秩 $r = \text{rank}(A) < \min\{m, n\}$ 时, 由于奇异值 $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_h = 0$, $h = \min\{m, n\}$, 故奇异值分解公式 (6.2.3) 可以简化为

$$A = U_r \Sigma_r V_r^H \quad (6.2.15)$$

式中

$$U_r = [u_1, u_2, \dots, u_r], \quad V_r = [v_1, v_2, \dots, v_r], \quad \Sigma_r = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$$

式 (6.2.15) 称为矩阵 A 的截尾奇异值分解 (truncated SVD) 或薄奇异值分解 (thin SVD)。与之形成对照, 式 (6.2.3) 则称为全奇异值分解 (full SVD)。

(7) 如果矩阵 $A_{m \times n}$ 具有秩 r , 则

- ① $m \times m$ 酉矩阵 U 的前 r 列组成矩阵 A 的列空间的标准正交基。
- ② $n \times n$ 酉矩阵 V 的前 r 列组成矩阵 A 的行空间 (或 A^H 的列空间) 的标准正交基。
- ③ V 的后 $n-r$ 列组成矩阵 A 的零空间的标准正交基。
- ④ U 的后 $m-r$ 列组成矩阵 A^H 的零空间的标准正交基。

顾名思义, 矩阵 A 的奇异值应该能够描述 A 的奇异性质。下面的定理从数学上严格地叙述了这一事实。

定理 6.2.2 ^[184] 令 $A \in C^{m \times n}$ ($m > n$) 的奇异值为

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

则

$$\sigma_k = \min_{E \in C^{m \times n}} \{\|E\|_F : \text{rank}(A + E) \leq k-1\}, \quad k = 1, 2, \dots, n \quad (6.2.16)$$

并且存在一满足 $\|E_k\|_F = \sigma_k$ 的误差矩阵 E 使得

$$\text{rank}(A + E_k) = k-1, \quad k = 1, 2, \dots, n$$

定理 6.2.2 表明, 奇异值与使得原矩阵 A 的秩减小 1 的误差矩阵 E_k 的 Frobenius 范数相等。如果原 $n \times n$ 矩阵 A 是正方的, 并且具有一个零奇异值, 则定理 6.2.2 表明, 该矩阵的秩减小 1 的误差矩阵 E 的 Frobenius 范数等于零。这意味着, 误差矩阵必然是一个零矩阵。换句话说, 根据定理 6.2.2, 当原 $n \times n$ 矩阵 A 有一个零奇异值时, 该矩阵的秩 $\text{rank}(A) \leq n-1$, 即原矩阵 A 本来就不是满秩的。因此, 如果一个正方形矩阵具有零奇异值, 则该矩阵必定是奇异矩阵。从这个角度讲, 零奇异值刻画了矩阵 A 的奇异性质。一个正方形矩阵只要有一个奇异值接近零, 那么这个矩阵就接近于奇异矩阵。推而广之, 一个非正方的矩阵如果有奇异值为零, 则说明这个长方矩阵一定不是列满秩的或者行满秩的。这种情况称为矩阵的秩亏缺, 它相对于矩阵的满秩是一种奇异现象。总之, 无论是正方还是长方矩阵, 零奇异值都刻画矩阵的奇异性。这就是矩阵奇异值的内在涵义。

下面以矩阵方程 (6.1.2) 的求解为例, 考查奇异值分解的几何意义。

首先, 我们可以把

$$\tilde{x} = V^H x \quad \text{或} \quad x = V \tilde{x} \quad (6.2.17)$$

看作是利用 V 进行的一种正交变换 (也可认为是一种旋转), 将 x 的各点旋转为 \tilde{x} 的各点。同样地, 我们也可以利用 U^H 对 b 作正交变换:

$$\tilde{b} = U^H b \quad (6.2.18)$$

即将 b 的各点旋转一定角度后变为 \tilde{b} 上的各点。现在, 将奇异值分解式 (6.2.3) 代入式 (6.1.2), 并利用式 (6.2.17) 和式 (6.2.18), 可得到

$$\tilde{b} = \Sigma \tilde{x} \longrightarrow \tilde{x} = \Sigma^\dagger \tilde{b}$$

于是, 线性方程组 (6.1.2) 的求解过程可以解释为一系列的线性变换操作, 即

$$b \xrightarrow{U} U^H b = \tilde{b} \xrightarrow{\Sigma} \Sigma^\dagger \tilde{b} = \tilde{x} \xrightarrow{V} V \tilde{x} = x$$

注意, Σ 的广义逆矩阵 Σ^\dagger 可直接计算为

$$\Sigma^\dagger = \begin{bmatrix} \Sigma^{-1} & O \\ O & O \end{bmatrix} \quad (6.2.19)$$

其中

$$\Sigma^{-1} = \text{diag}(1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_r) \quad (6.2.20)$$

把 $m \times n$ 矩阵 A 视作从 n 维 (复数) 向量空间 C^n 到 m 维 (复数) 向量空间 C^m 的线性映射有时是很方便的。此时, 关于奇异值分解的唯一性, 有以下结果 [518]。

- (1) 非零奇异值的个数 r 和它们的值 $\sigma_1, \sigma_2, \dots, \sigma_r$ 相对于矩阵 A 是唯一确定的。
- (2) 若 $\text{rank}(A) = r$, 则满足 $Ax = 0$ 的 $x (\in C^n)$ 的集合即 A 的零空间 $\text{Null } A (\subseteq C^n)$ 是 $n-r$ 维的, 因此可选择正交基 $\{v_{r+1}, v_{r+2}, \dots, v_n\}$ 作为 A 在 C^n 内的零空间。从这个意义上讲, V 的列向量张成的 C^n 的子空间 $\text{Null}(A)$ 是唯一确定的, 但是各个向量只要能组成该子空间的正交基, 它们就可以自由地选择。
- (3) 可以表示成 $y = Ax$ 的 $y (\in C^m)$ 的集合组成 A 的像空间 $\text{Im}A$, 它是 r 维的, 而 $\text{Im}A$ 的正交补空间 $(\text{Im}A)^\perp$ 是 $m-r$ 维的, 因此可选择 $\{u_{r+1}, u_{r+2}, \dots, u_m\}$ 作为 $\text{Im}A$ 在 C^m 内的正交补空间内的正交基。由 U 的列向量 $u_{r+1}, u_{r+2}, \dots, u_m$ 张成的 C^m 的子空间 $(\text{Im}A)^\perp$ 是唯一确定的。
- (4) 若 σ_i 是单奇异值 (即 $\sigma_i \neq \sigma_j, \forall j \neq i$), 则 v_i 和 u_i 除相差一相角 (A 为实数矩阵时, 相差一符号) 外是唯一确定的。也就是说, v_i 和 u_i 同时乘以 $e^{j\theta}$ ($j = \sqrt{-1}$, 且 θ 为实数) 后, 它们仍然分别是矩阵 A 的右和左奇异向量。

6.2.2 奇异值的性质

下面分几种情况, 分别详细讨论奇异值的各种性质。为统一计, 令矩阵 A 和 B 均为 $m \times n$ 矩阵, 并且 $r_A = \text{rank}(A)$, $p = \min\{m, n\}$ 。

设矩阵 A 的奇异值排列为

$$\sigma_{\max} = \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{p-1} \geq \sigma_p = \sigma_{\min} \geq 0 \quad (6.2.21)$$

并且用 $\sigma_i(B)$ 表示矩阵 B 的第 i 个大奇异值。

矩阵的各种变形与奇异值的变化有以下关系。

(1) $m \times n$ 矩阵 A 的共轭转置 A^H 的奇异值分解为

$$A^H = V \Sigma^T U^H \quad (6.2.22)$$

即矩阵 A 和 A^H 具有完全相同的奇异值。

(2) P 和 Q 分别为 $m \times m$ 和 $n \times n$ 酉矩阵时, PAQ^H 的奇异值分解由

$$PAQ^H = \tilde{U} \Sigma \tilde{V}^H \quad (6.2.23)$$

给出, 其中, $\tilde{U} = PU$, $\tilde{V} = QV$ 。就是说, 矩阵 PAQ^H 与 A 具有相同的奇异值, 即奇异值具有酉不变性, 但奇异向量不同。

(3) $A^H A$, AA^H 的奇异值分解分别为

$$A^H A = V \Sigma^T \Sigma V^H, \quad AA^H = U \Sigma \Sigma^T U^H \quad (6.2.24)$$

其中

$$\Sigma^T \Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, \overbrace{0, \dots, 0}^{n-r}) \quad (6.2.25)$$

$$\Sigma \Sigma^T = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, \overbrace{0, \dots, 0}^{m-r}) \quad (6.2.26)$$

注: $A^H A$ 和 AA^H 均为 Hermitian 矩阵。Hermitian 矩阵的奇异值分解与特征值分解是一致的。

(4) $m \times n$ 矩阵 A 的奇异值分解与 $n \times m$ 维 Moore-Penrose 广义逆矩阵 A^\dagger 之间存在下列关系:

$$A^\dagger = V \Sigma^\dagger U^H \quad (6.2.27)$$

其中, Σ^\dagger 由式 (6.2.19) 给定。

证明 前三个性质显然。为了证明性质 (4), 令 $G = V \Sigma^\dagger U^H$ 。于是, 有

$$\begin{aligned} AGA &= (U \Sigma V^H)(V \Sigma^\dagger U^H)(U \Sigma V^H) \\ &= U \Sigma \Sigma^\dagger \Sigma V^H = U \Sigma V^H = A \\ GAG &= (V \Sigma^\dagger U^H)(U \Sigma V^H)(V \Sigma^\dagger U^H) \\ &= V \Sigma^\dagger \Sigma \Sigma^\dagger U^H = V \Sigma^\dagger U^H = G \\ AG &= U \Sigma \Sigma^\dagger U^H = (AG)^H \\ GA &= V \Sigma^\dagger \Sigma V^H = (GA)^H \end{aligned}$$

即 G 满足 Moore-Penrose 广义逆矩阵的定义。因此, G 是矩阵 A 的 Moore-Penrose 广义逆矩阵, 即性质 (4) 成立。■

虽然 U 和 V 相对于 A 不是唯一确定的, 但广义逆矩阵 A^\dagger 是唯一确定的。特别地, 若 A 是一个正方的非奇异矩阵, 则 $A^\dagger = A^{-1}$ 。因此, 在这一情况下, 如果 A 的奇异值是 $\sigma_1, \sigma_2, \dots, \sigma_n$, 那么 A^{-1} 的奇异值就是 $1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_n$ 。

关于矩阵和它的子矩阵的奇异值之间的关系, 有下面的定理, 常被称为奇异值交织定理 (interlacing theorem for singular values)。

定理 6.2.3^{[225, Eq. (3.1.3)], [233]} 令 A 是一个 $m \times n$ 矩阵, 其奇异值 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$, 其中, $r = \min\{m, n\}$ 。若 $p \times q$ 矩阵 B 是 A 的子矩阵, 其奇异值 $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_{\min\{p, q\}}$, 则

$$\sigma_i \geq \gamma_i, \quad i = 1, 2, \cdots, \min\{p, q\} \quad (6.2.28)$$

并且

$$\gamma_i \geq \sigma_{i+(m-p)+(n-q)}, \quad i \leq \min\{p+q-m, p+q-n\} \quad (6.2.29)$$

矩阵的奇异值与矩阵的范数、行列式、条件数、特征值等有着密切的关系。

1. 奇异值与范数的关系

矩阵 A 的谱范数等于 A 的最大奇异值, 即

$$\|A\|_{\text{spec}} = \sigma_1 \quad (6.2.30)$$

根据矩阵的奇异值分解定理, 并注意到矩阵 A 的 Frobenius 范数 $\|A\|_F$ 是酉不变的, 即 $\|U^H A V\|_F = \|A\|_F$, 故有

$$\|A\|_F = \left[\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right]^{1/2} \quad (6.2.31)$$

$$\begin{aligned} &= \|U^H A V\|_F = \|\Sigma\|_F \\ &= \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2} \end{aligned} \quad (6.2.32)$$

即是说, 任何一个矩阵的 Frobenius 范数等于该矩阵所有非零奇异值平方和的正平方根。

考虑矩阵 A 的秩 k 近似, 并将其记作 A_k , 其中, $k < r = \text{rank}(A)$ 。矩阵 A_k 定义如下:

$$A_k = \sum_{i=1}^k \sigma_i u_i v_i^H, \quad k < r$$

则 A 与秩为 k 的任一矩阵 B 之差的 l_1 和 Frobenius 范数分别为

$$\min_{\text{rank}(B)=k} \|A - B\|_1 = \|A - A_k\|_1 = \sigma_{k+1} \quad (6.2.33)$$

和

$$\min_{\text{rank}(B)=k} \|A - B\|_F^2 = \|A - A_k\|_F^2 = \sigma_{k+1}^2 + \sigma_{k+2}^2 + \cdots + \sigma_r^2 \quad (6.2.34)$$

这一重要结果是许多概念和应用的基础。例如, 总体最小二乘、数据压缩、图像增强、动态系统实现理论, 以及线性方程的求解等问题都需要用一个低秩矩阵近似 A 。

2. 奇异值与行列式的关系

设 \mathbf{A} 是 $n \times n$ 正方矩阵。由于酉矩阵的行列式之绝对值等于 1, 所以由定理 6.2.1 有

$$|\det(\mathbf{A})| = |\det \Sigma| = \sigma_1 \sigma_2 \cdots \sigma_n \quad (6.2.35)$$

若所有 σ_i 都不等于零, 则 $|\det(\mathbf{A})| \neq 0$, 这表明 \mathbf{A} 是非奇异的。如果至少有一个 $\sigma_i (i > r)$ 等于零, 便有 $\det(\mathbf{A}) = 0$, 即 \mathbf{A} 是奇异的。这就是之所以把全部 σ_i 值统称为奇异值的原因。综合式 (6.2.34) 和式 (6.2.35), 对于一个 $n \times n$ 矩阵 \mathbf{A} , 下列不等式成立:

$$\left. \begin{aligned} n\sigma_1 &\geq \|\mathbf{A}\|_F \geq \sigma_1 \\ \sigma_1^n &\geq \sigma_1^{n-1}\sigma_n \geq |\det(\mathbf{A})| \geq \sigma_n^n \\ \|\mathbf{A}\|_F &\geq \sigma_1 \geq |\det(\mathbf{A})|^{1/n} \\ |\det(\mathbf{A})|^{1/n} &\geq \sigma_n \geq |\det(\mathbf{A})|/\|\mathbf{A}\|_F^{n-1} \\ \|\mathbf{A}\|_F^n/|\det(\mathbf{A})| &\geq \sigma_1/\sigma_n \geq \max\left\{1, \frac{1}{n}\|\mathbf{A}\|_F/|\det(\mathbf{A})|^{1/n}\right\} \end{aligned} \right\} \quad (6.2.36)$$

这些不等式虽然是粗略的评价, 但有时是有用的。

3. 奇异值与条件数的关系

对于一个 $m \times n$ 矩阵 \mathbf{A} , 其条件数也可以利用奇异值定义为

$$\text{cond}(\mathbf{A}) = \sigma_1/\sigma_p, \quad p = \min\{m, n\} \quad (6.2.37)$$

由定义式 (6.2.37) 可以看出, 条件数是一个大于或等于 1 的正数, 因为 $\sigma_1 \geq \sigma_p$ 。显然, 由于至少有一个奇异值 $\sigma_p = 0$, 故奇异矩阵的条件数为无穷大, 而条件数虽然不是无穷大, 但却很大时, 就称 \mathbf{A} 是接近奇异的。这意味着, 当条件数很大时, \mathbf{A} 的行向量或列向量的线性相关性很强。另由定义式 (6.1.8) 易知, 正交或酉矩阵 \mathbf{V} 的条件数等于 1。从这个意义上讲, 正交或酉矩阵是“理想条件”的。式 (6.2.37) 也可用作条件数 $\text{cond}(\mathbf{A})$ 的评价。

考虑超定方程 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 。此时, 由于 $\mathbf{A}^H\mathbf{A}$ 的奇异值分解为

$$\mathbf{A}^H\mathbf{A} = \mathbf{V}\Sigma^2\mathbf{V}^H \quad (6.2.38)$$

即矩阵 $\mathbf{A}^H\mathbf{A}$ 的最大和最小奇异值分别是矩阵 \mathbf{A} 的最大和最小奇异值的平方, 故

$$\text{cond}(\mathbf{A}^H\mathbf{A}) = \frac{\sigma_1^2}{\sigma_n^2} = [\text{cond}(\mathbf{A})]^2 \quad (6.2.39)$$

换言之, 矩阵 $\mathbf{A}^H\mathbf{A}$ 的条件数是矩阵 \mathbf{A} 的条件数的平方倍。

4. 奇异值与特征值的关系

设 $n \times n$ 正方对称矩阵 \mathbf{A} 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n (|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|)$, 奇异值为 $\sigma_1, \sigma_2, \dots, \sigma_n (\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0)$, 则 $\sigma_i \geq |\lambda_i| \geq \sigma_n (i = 1, 2, \dots, n)$, $\text{cond}(\mathbf{A}) \geq |\lambda_1|/|\lambda_n|$ 。

特别值得指出的是, 奇异值分解提供了在实际信号处理中发生的一些重要问题的定量答案。

问题 1: 一矩阵与低秩矩阵是如何接近的?

答案: 如果 $A \in C^{m \times n}$ 具有奇异值

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0$$

则 $\text{rank}(A) = r$, 且

$$\min_{\text{rank}(B) \leq r} \|A - B\|_F^2 = \sigma_{r+1}^2$$

特别地, 若 $A \in C^{n \times n}$ 非奇异, 则 σ_n 是到奇异矩阵集合的距离。因此, $1/\text{cond}(A) = \sigma_n/\sigma_1 = \sigma_n/\|A\|_F$ 是 A 接近奇异的相对测度。

问题 2: 一矩阵的值域与零空间是什么?

答案: 如果 $U = [u_1, u_2, \cdots, u_m]$ 和 $V = [v_1, v_2, \cdots, v_n]$ 分别是左和右奇异向量矩阵的列分块, 并且 $\text{rank}(A) = r$, 则 $\text{Null}(A) = \text{Span}\{v_{r+1}, v_{r+2}, \cdots, v_n\}$, $\text{Range}(A) = \text{Span}\{u_1, u_2, \cdots, u_r\}$ 。即是说, 一矩阵的值域是与非零奇异值对应的奇异向量张成的子空间, 而零空间则是与零奇异值对应的奇异向量张成的子空间。

问题 3: 怎样度量 C^m 中两个 k 维子空间是否接近?

答案: 假定 $n \times k$ 矩阵 $Y = [y_1, y_2, \cdots, y_k]$ 和 $Z = [z_1, z_2, \cdots, z_k]$ 具有正交列, 并且 $S_1 = \text{Span}(Y)$ 和 $S_2 = \text{Span}(Z)$ 分别是矩阵 Y 和 Z 的列向量张成的子空间。若 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k$ 是 $Y^H Z$ 的奇异值, 则

$$\text{dist}(S_1, S_2) = \min_{y \in S_1, z \in S_2, \|z\|_2=1} \|y - z\|_2 = \sqrt{1 - \sigma_1^2}$$

上述问题答案的证明可在文献 [184] 中找到。

6.2.3 奇异值的性质汇总

为了方便读者参考, 下面汇总了矩阵的奇异值的性质。

1. 奇异值服从的等式关系 [290]

- (1) 矩阵 $A_{m \times n}$ 和其复共轭转置矩阵 A^H 具有相同的奇异值。
- (2) 矩阵 $A_{m \times n}$ 的非零奇异值是 AA^H 或者 $A^H A$ 的非零特征值的正平方根。
- (3) $\sigma > 0$ 是矩阵 $A_{m \times n}$ 的单奇异值, 当且仅当 σ^2 是 AA^H 或 $A^H A$ 的单特征值。
- (4) 若 $p = \min\{m, n\}$, 且 $\sigma_1, \sigma_2, \cdots, \sigma_p$ 是矩阵 $A_{m \times n}$ 的奇异值, 则

$$\text{tr}(A^H A) = \sum_{i=1}^p \sigma_i^2$$

- (5) 矩阵行列式的绝对值等于矩阵奇异值之乘积, 即 $|\det(A)| = \sigma_1 \sigma_2 \cdots \sigma_n$ 。

- (6) 矩阵 A 的谱范数等于 A 的最大奇异值, 即 $\|A\|_{\text{spec}} = \sigma_{\max}$ 。

(7) 若 $m \geq n$, 则对于矩阵 $A_{m \times n}$, 有

$$\begin{aligned}\sigma_{\min}(A) &= \min \left\{ \left(\frac{x^H A^H A x}{x^H x} \right)^{1/2} : x \neq 0 \right\} \\ &= \min \left\{ (x^H A^H A x)^{1/2} : x^H x = 1, x \in C^n \right\}\end{aligned}$$

(8) 若 $m \geq n$, 则对于矩阵 $A_{m \times n}$, 有

$$\begin{aligned}\sigma_{\max}(A) &= \max \left\{ \left(\frac{x^H A^H A x}{x^H x} \right)^{1/2} : x \neq 0 \right\} \\ &= \max \left\{ (x^H A^H A x)^{1/2} : x^H x = 1, x \in C^n \right\}\end{aligned}$$

(9) 若 $m \times m$ 矩阵 A 非奇异, 则

$$\frac{1}{\sigma_{\min}(A)} = \max \left\{ \left(\frac{x^H (A^{-1})^H A^{-1} x}{x^H x} \right)^{1/2} : x \neq 0, x \in C^n \right\}$$

(10) 若 $A = U \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix} V^H$ 是 $m \times n$ 矩阵 A 的奇异值分解, 则 A 的 Moore-Penrose 逆矩阵

$$A^\dagger = V \begin{bmatrix} \Sigma_1^{-1} & O \\ O & O \end{bmatrix} U^H$$

(11) 若 $\sigma_1, \sigma_2, \dots, \sigma_p$ 是 $m \times n$ 矩阵 A 的非零奇异值 (其中, $p = \min\{m, n\}$), 则矩阵 $\begin{bmatrix} O & A \\ A^H & O \end{bmatrix}$ 具有 $2p$ 个非零奇异值 $\sigma_1, \dots, \sigma_p, -\sigma_1, \dots, -\sigma_p$ 和 $|m-n|$ 个零奇异值。

2. 奇异值服从的不等式关系 [224],[225],[270],[81],[290]

(1) 若 A 和 B 是 $m \times n$ 矩阵, 则对于 $1 \leq i, j \leq p, i+j \leq p+1$ ($p = \min\{m, n\}$), 有

$$\sigma_{i+j-1}(A+B) \leq \sigma_i(A) + \sigma_j(B)$$

特别地, 当 $j=1$ 时, $\sigma_i(A+B) \leq \sigma_i(A) + \sigma_1(B)$, $i=1, 2, \dots, p$ 成立。

(2) 对矩阵 $A_{m \times n}, B_{m \times n}$, 有

$$\sigma_{\max}(A+B) \leq \sigma_{\max}(A) + \sigma_{\max}(B)$$

(3) 若 A 和 B 是 $m \times n$ 矩阵, 则

$$\sum_{j=1}^p [\sigma_j(A+B) - \sigma_j(A)]^2 \leq \|B\|_F^2, \quad p = \min\{m, n\}$$

(4) 若 $A_{m \times m} = [a_1, a_2, \dots, a_m]$ 的奇异值 $\sigma_1(A) \geq \sigma_2(A) \geq \dots \geq \sigma_m(A)$, 则

$$\sum_{j=1}^k [\sigma_{m-k+j}(A)]^2 \leq \sum_{j=1}^k a_j^H a_j \leq \sum_{j=1}^k [\sigma_j(A)]^2, \quad k = 1, 2, \dots, m$$

(5) 若 $p = \min\{m, n\}$, 且 $A_{m \times n}$ 和 $B_{m \times n}$ 的奇异值排列为 $\sigma_1(A) \geq \sigma_2(A) \geq \dots \geq \sigma_p(A)$, $\sigma_1(B) \geq \sigma_2(B) \geq \dots \geq \sigma_p(B)$ 和 $\sigma_1(A+B) \geq \sigma_2(A+B) \geq \dots \geq \sigma_p(A+B)$, 则

$$\sigma_{i+j-1}(AB^H) \leq \sigma_i(A)\sigma_j(B), \quad 1 \leq i, j \leq p, i+j \leq p+1$$

(6) 设 $m \times (n-1)$ 矩阵 B 是删去 $m \times n$ 矩阵 A 任意一列得到的矩阵, 并且它们的奇异值都按照非降顺序排列, 则

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \sigma_2(B) \geq \dots \geq \sigma_h(A) \geq \sigma_h(B) \geq 0$$

式中, $h = \min\{m, n-1\}$ 。

(7) 设 $(m-1) \times n$ 矩阵 B 是删去 $m \times n$ 矩阵 A 任意一行得到的矩阵, 并且它们的奇异值都按照非降顺序排列, 则

$$\sigma_1(A) \geq \sigma_1(B) \geq \sigma_2(A) \geq \sigma_2(B) \geq \dots \geq \sigma_h(A) \geq \sigma_h(B) \geq 0$$

式中, $h = \min\{m, n-1\}$ 。

(8) 矩阵 $A_{m \times n}$ 的最大奇异值满足不等式

$$\sigma_{\max}(A) \geq \left[\frac{1}{n} \text{tr}(A^H A) \right]^{1/2}$$

6.2.4 秩亏缺最小二乘解

在奇异值分析的应用中, 常常需要用一个低秩的矩阵逼近一个含噪声或扰动的矩阵。下面的定理给出了逼近质量的评价。

定理 6.2.4 令 $A \in R^{m \times n}$ 的奇异值分解由 $A = \sum_{i=1}^p \sigma_i u_i v_i^T$ 给出, 其中, $p = \text{rank}(A)$ 。

若 $k < p$, 并且 $A_k = \sum_{i=1}^k \sigma_i u_i v_i^T$, 则逼近质量可分别使用谱范数和 Frobenius 范数度量:

$$\min_{\text{rank}(B)=k} \|A - B\|_{\text{spec}} = \|A - A_k\|_{\text{spec}} = \sigma_{k+1} \quad (6.2.40)$$

$$\min_{\text{rank}(B)=k} \|A - B\|_F = \|A - A_k\|_F = \sqrt{\sum_{i=k+1}^q \sigma_i^2} \quad (6.2.41)$$

式中, $q = \min\{m, n\}$ 。

证明 详见文献 [140], [315] 或 [233]。

在信号处理和系统理论中, 最常见的线性方程组 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 是超定的和非满秩即秩亏缺的, 也就是说, 矩阵 $\mathbf{A} \in \mathbb{C}^{m \times n}$ 的行数 m 比列数 n 大, 且 $r = \text{rank}(\mathbf{A}) < n$ 。令 \mathbf{A} 的奇异值分解由式 $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ 给出, 其中, $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$ 。考察

$$\mathbf{G} = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^H \quad (6.2.42)$$

式中, $\mathbf{\Sigma}^\dagger = \text{diag}(1/\sigma_1, \dots, 1/\sigma_r, 0, \dots, 0)$ 。由奇异值的性质 (4) 知, \mathbf{G} 是 \mathbf{A} 的 Moore-Penrose 广义逆矩阵。因此,

$$\hat{\mathbf{x}} = \mathbf{G}\mathbf{b} = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^H\mathbf{b} \quad (6.2.43)$$

给出最小二乘最小范数解。此时, 解 $\hat{\mathbf{x}}$ 的误差矩阵由

$$\mathbf{\Sigma}_{\hat{\mathbf{x}}} = (\mathbf{A}^H\mathbf{A})^{-1} = \mathbf{V}(\mathbf{\Sigma}^T\mathbf{\Sigma})^\dagger\mathbf{V}^H \quad (6.2.44)$$

给出。式 (6.2.43) 可表示为

$$\mathbf{x}_{\text{LS}} = \sum_{i=1}^r (\mathbf{u}_i^H \mathbf{b} / \sigma_i) \mathbf{v}_i$$

它是最小二乘问题

$$\min \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 \quad (6.2.45)$$

的最小范数解, 相应的最小残差为

$$\rho_{\text{LS}} = \|\mathbf{A}\mathbf{x}_{\text{LS}} - \mathbf{b}\|_2 = \|[\mathbf{u}_{r+1}, \dots, \mathbf{u}_m]^H \mathbf{b}\|_2 \quad (6.2.46)$$

应用奇异值分解求解最小二乘问题的方法常简称为奇异值分解方法。虽然在理论上, 当 $i > r$ 时奇异值 $\sigma_i = 0$, 但是计算出来的奇异值 $\hat{\sigma}_i$, $i > r$ 并不会等于零, 有时甚至表现出比较大的扰动。因此, 需要有计算秩 r 的估计值 \hat{r} 的方法。在信号处理和系统理论中, 常将该估计值称为“有效秩”。

有效秩确定有以下两种常用方法。

1. 归一化奇异值方法

计算归一化奇异值

$$\bar{\sigma}_i = \frac{\hat{\sigma}_i}{\hat{\sigma}_1} \quad (6.2.47)$$

选择满足准则

$$\bar{\sigma}_i \geq \epsilon \quad (6.2.48)$$

的最大整数作为有效秩的估计值 \hat{r} 。显然, 这一准则等价于选择满足

$$\hat{\sigma}_i \geq \epsilon \cdot \hat{\sigma}_1 \quad (6.2.49)$$

的最大整数 \hat{r} 。式中, ϵ 是某个很小的正数, 它根据计算机精度与 (或) 数据精度选取。例如, 选取 $\epsilon = 0.1$ 或者 $\epsilon = 0.05$ 等。

2. 范数比方法

令 $m \times n$ 矩阵 A_k 是原 $m \times n$ 矩阵 A 的秩 k 近似, 定义该近似矩阵与原矩阵的 Frobenius 范数比为

$$\nu(k) = \frac{\|A_k\|_F}{\|A\|_F} = \frac{\sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_k^2}}{\sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_h^2}}, \quad h = \min\{m, n\} \quad (6.2.50)$$

并选择满足

$$\nu(k) \geq \alpha \quad (6.2.51)$$

的最大整数作为有效秩估计 \hat{r} , 其中, α 是接近于 1 的阈值, 例如 $\alpha = 0.997$ 等。

采用以上两种准则确定出有效秩 \hat{r} 后, 可将

$$\hat{x}_{LS} = \sum_{i=1}^{\hat{r}} (\hat{u}_i^H b / \hat{\sigma}_i) \hat{v}_i \quad (6.2.52)$$

看作是真实最小二乘解 x_{LS} 的一个合理近似。显而易见, 这种解就是方程组 $A_{\hat{r}} x = b$ 的最小二乘解, 其中

$$A_{\hat{r}} = \sum_{i=1}^{\hat{r}} \sigma_i u_i v_i^H \quad (6.2.53)$$

在最小二乘问题中, 用 $A_{\hat{r}}$ 代替 A 相当于过滤掉小的奇异值。当 A 是从有噪声的观测数据得到时, 这种过滤能够起很大的作用。容易观察到, 式 (6.2.52) 给出的最小二乘解 \hat{x}_{LS} 仍然包含了 n 个参数。然而, 由于线性方程 $Ax = b$ 秩亏缺意味着 x 中只有 r 个参数是独立的, 其他参数是这 r 个独立参数的重复作用或线性相关的结果。在许多应用中, 当然希望能够求出这 r 个线性无关的参数, 而不是包含了冗余因素的 n 个参数。换言之, 我们的目的是只估计主要因素, 并剔除掉次要因素。在线性代数中, 这相当于如何从矩阵 A 的 n 列中挑选出 r 个线性无关的列来。怎样挑选这些线性无关的列称为子集选择问题。

在秩亏缺的情况下, 利用 SVD 进行子集选择是必要的。Golub 等人曾提出过一种基于 SVD 的子集选择方法 [183]。

算法 6.2.1 (子集选择算法)

步骤 1 计算 A 的 SVD, 并确定 A 的有效秩 \hat{r} 。

步骤 2 计算置换矩阵 P , 使得在 $AP = [B_1, B_2]$ 中的矩阵 $B_1 \in C^{m \times \hat{r}}$ 的列是“足够线性无关的”。

步骤 3 将 $Ax = 0$ 的 LS 问题变换为求 $AP \begin{bmatrix} z \\ 0 \end{bmatrix} = b$ 的 LS 解 $z \in C^{\hat{r}}$, 即求 $\|B_1 z - b\|_2$ 的极小化变量 z 。

由于

$$\min_{z \in C^{\hat{r}}} \|B_1 z - b\|_2 \geq \|Ax - b\|_2$$

所以置换矩阵 P 应选得使残差 $(I - B_1 B_1^\dagger)b$ 的范数尽可能小。但是, 这有可能产生不稳定的解。例如, 对于

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1+\epsilon & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

若取 $\hat{r} = 2$ 和 $P = I$, 则 $\min \|B_1 z - b\|_2 = 0$, 但 $\|B_1^\dagger b\|_2 = O(1/\epsilon)$, 即 LS 解 $z = B_1^\dagger b$ 不稳定。这个例子说明, 应该在选择列的线性无关性和它给出的残差范数 (即解的稳定性) 之间取折衷。考虑到这些因素, Golub 等人利用基于 SVD 的子集选择, 提出了求方程 $Ax = b$ 中 x 的独立参数的最小二乘解的下述实际算法^[183], 称之为低秩 LS 方法。

算法 6.2.2 (低秩 LS 方法) 给定 $A \in R^{m \times n}$, $b \in R^m$, 本算法计算

$$x = \arg \min_z \left\| AP \begin{bmatrix} z \\ 0 \end{bmatrix} - b \right\|_2$$

步骤 1 计算 SVD 即 $U^T A V = \text{diag}(\sigma_1, \dots, \sigma_n)$, 确定有效秩 \hat{r} , 并根据 \hat{r} 将 V 分块为

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

其中, $V_{11} \in R^{\hat{r} \times \hat{r}}$, 然后存储 V_{11} 和 V_{21} 。

步骤 2 利用列主元 QR 算法 (算法 6.2.3) 计算 $Q^T [V_{11}^T, V_{21}^T] P = [R_{11}, R_{12}]$, 然后计算 $AP = [B_1, B_2]$, 其中, $B_1 \in R^{m \times \hat{r}}$ 。

步骤 3 计算 $z = (B_1^T B_1)^{-1} B_1^T b$ 。

下面是步骤 2 用到的列主元 QR 算法。

算法 6.2.3 (列主元 QR 分解)^[184] 给定矩阵 $A \in R^{m \times n}$, 其中, $m \geq n$ 。下面的算法计算 $r = \text{rank}(A)$ 和列主元 QR 分解

$$Q^T A P = \begin{bmatrix} R_{11} & R_{12} \\ O & O \end{bmatrix}$$

其中, $R_{11} \in R^{r \times r}$ 是上三角的非奇异矩阵。作为输出结果, A 的上三角部分存放 R 的上三角部分, 置换矩阵 P 用整数向量 piv 编码 (若 piv 的第 j 个元素等于整数 m , 则 P 的第 j 列仅第 m 个元素为 1, 而其他元素皆为零)。

for $j = 1, 2, \dots, n$

$c(j) = A(1:m, j)^T A(1:m, j)$

end

$r = 0; \tau = \max\{c(j), \dots, c(n)\}$

求满足 $c(k) = \tau$ 的最小整数 k ($1 \leq k \leq n$)

while $\tau > 0$

$r = r + 1$

$piv(r) = k; A(1:m, r) \leftrightarrow A(1:m, k); c(\tau) \leftrightarrow c(k)$

$v(r:m) = \text{house}(A(r:m, r))$

$A(r:m, r:n) = \text{row.house}(A(r:m, r:n), v(r:m))$

```

 $A(r+1:m, r) = v(r+1:m)$ 
for  $i = r+1:n$ 
     $c(i) = c(i) - A(r, i)^2$ 
end
if  $r < n$ 
     $\tau = \max\{c(r+1), \dots, c(n)\}$ 
    求满足  $c(k) = \tau$  的最小整数  $k (r+1 \leq k \leq n)$ 
else
     $\tau = 0$ 
end
end

```

例 6.2.1^[184] 假定

$$A = \begin{bmatrix} 3 & 4 & -1 \\ 7 & 4 & -3 \\ 2 & 5 & 3 \\ -1 & 4 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

得到 $\text{rank}(A) = 2$ 和

$$x_{LS} = \begin{bmatrix} 0.0815 \\ 0.1545 \\ 0.0730 \end{bmatrix}$$

应用算法 6.2.2, 则得到

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} 0.0845 \\ 0.2275 \\ 0.0000 \end{bmatrix}$$

顺便指出, $\|b - Ax_{LS}\|_2 \approx \|b - Ax\|_2 = 0.1966$ 。

有必要指出, 上述最小二乘解只考虑了线性方程一边的误差, 它的数值性能比不上同时考虑方程两边误差的总体最小二乘方法 (详见第 7 章)。与低秩最小二乘方法一样, 也有低秩总体最小二乘方法, 但是它们的子集选择方法明显不同。

6.3 奇异值分解的数值计算

前面介绍了奇异值分解的定义、性质及求解线性方程的奇异值分解方法。本节讨论奇异值分解 $A = U\Sigma V^T$ 的数值计算: 给定矩阵 $A_{m \times n}$, 如何用尽可能高的精度计算其奇异值, 右奇异向量矩阵 $V_{n \times n}$ 与 (或) 左奇异向量矩阵 $U_{m \times m}$ 。

求解一般矩阵的奇异值问题的最常用算法可以分为两大类:

- (1) QR 分解
- (2) Jacobi 旋转

6.3.1 奇异值分解的 QR 分解算法

这类算法通常分两个阶段进行：第一阶段为矩阵的二重对角化，即应用 Householder 变换将矩阵 $A_{m \times n}$ 变换为二重对角矩阵（除主对角线及其上面一条对角线的元素外，其他元素全为零）。第二阶段为 QR 分解，即保持二重对角矩阵的形式不变，利用正交变换使上一条对角线的元素逐渐减小，使矩阵接近对角矩阵。此时，需要具体构造正交矩阵 V ，但 U 可以不予保留。

应用 Householder 变换进行二重对角化又分为两个步骤：第一步使 A 的第 1 列除 $(1,1)$ 元素以外的其他元素全部变换为零；尔后，对除 $(1,1)$ 元素以外的第 1 行再作 Householder 变换，以便使 $(1,2)$ 元素以右的其他元素全部变换为零。第二步对除去第 1 行和第 1 列后剩下的其余部分进行与第一步类似的 Householder 变换，使得 $(1,2)$ 和 $(2,2)$ 以外的第 2 列元素和 $(2,2)$ 和 $(2,3)$ 以外的第 2 行元素全部变换为零。如此重复，直到经过 $(2n-2)$ 次 Householder 变换，实现矩阵 A 的二重对角化，如图 6.3.1 所示。与此同时，对每行进行的 Householder 变换也同时对 V 进行，其中， V 的初始值设为单位矩阵 I 。

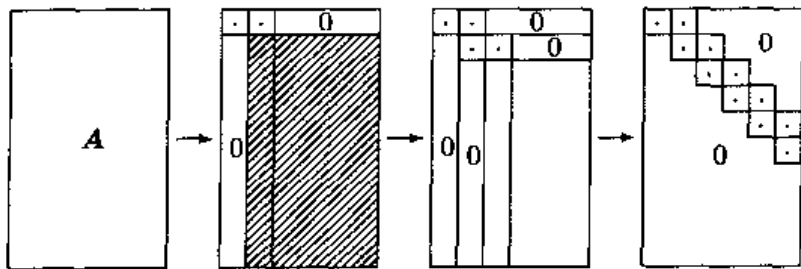


图6.3.1 利用Householder变换进行二重对角化

将对角元素记作 q_i ，上对角元素记作 e_i （定义 $e_1 = 0$ ），并舍去二重对角矩阵下面的零元素，从而剩下一个 $n \times n$ 的正方矩阵 G 。

在第二阶段即 QR 分解阶段，对 G 应用 QR 方法进行对角化，即反复应用正交变换，使上对角元素逐渐减小。利用 Givens 旋转矩阵 T_2 右乘矩阵 G 时， G 的第 1,2 列将发生变化，原先为零的 $(2,1)$ 元素会变为非零值。由于它是多余出来的，所以要用 Givens 矩阵 S_2^T 左乘，以使 $(2,1)$ 元素变作零。但是，由于 G_2^T 是左乘，第 1,2 行将发生变化，原来等于零的 $(1,3)$ 元素又会有非零值。然后，为了消去 $(1,3)$ 元素，我们又需要右乘 Givens 矩阵 T_3 。这样一来，又会产生非零的 $(3,2)$ 元素。重复这样的 Givens 旋转过程，最后为了消去 $(n, n-1)$ 元素而左乘 S_n^T 。由于没有 $(n+1)$ 列存在，所以最后仍然能够得到一个二重对角矩阵。上述过程称作追赶 (chasing) 过程。图 6.3.2 示出了追赶法消零的顺序。

在实际计算时，将矩阵 G 分块成尽可能小的二重对角矩阵可以加快追赶过程的收敛。因此，在利用 Givens 旋转进行追赶前，首先应分析全部的 q_i 和 e_i ，检查一下是否有等于零的元素。例如，如果 $e_i = 0$ ，则可以将 G 分块成两个独立的二重对角矩阵，即有

$$G = \begin{bmatrix} G_1 & O \\ O & G_2 \end{bmatrix}$$

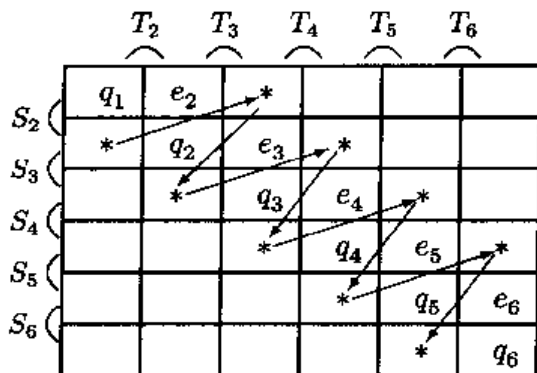


图6.3.2 追赶法消零的顺序

这样,即可对 G_1 和 G_2 分别进行独立的对角化。

需要强调的是,计算矩阵 A 的奇异值分解时,如果先通过正交变换把它变成三角矩阵形式,那么奇异值分解的数值稳定性通常将增加 [75]。

6.3.2 奇异值分解的精确计算

在应用计算机进行奇异值分解的数值运算时,通常使用浮点运算。令 $fl(\cdot)$ 表示某种运算的浮点计算结果,则有 [123]

$$\left. \begin{aligned} fl(a \pm b) &= a(1 + \varepsilon_1) + b(1 + \varepsilon_2) \\ fl(a \times b) &= (a \times b) \times (1 + \varepsilon_3) \\ fl(a/b) &= (a/b)(1 + \varepsilon_4) \\ fl(\sqrt{2}) &= \sqrt{2}(1 + \varepsilon_5) \end{aligned} \right\} \quad (6.3.1)$$

式中, $|\varepsilon_i| \leq \varepsilon$, 并且 $\varepsilon \ll 1$ 代表计算机精度。

由于初始矩阵 A 的元素往往是通过观测或者计算得到的,而观测数据或者计算不可避免地存在一定的误差,因此初始矩阵元素本身存在小的相对不确定度。显然,即使对这样的矩阵使用完全精确的奇异值分解计算方法,所得到的奇异值和奇异向量元素也存在固有的不确定度。若使用某种方法计算得到的奇异值和奇异向量元素的误差界接近它们固有的不确定度,则称这种方法是求解奇异值问题的精确算法,并称这种算法得到的奇异值和奇异向量的计算结果是精确的。基于这一考虑,Demmel 与 Veselic 于 1992 年提出了一种在一般矩阵 $A_{m \times n}$ 的右边使用单边 Jacobi 旋转 (称为右边 Jacobi 旋转) 的浮点运算,求解奇异值问题的下列算法 [123]。

算法 6.3.1 (奇异值分解的右边 Jacobi 算法) 令 tol 是用户自定义的迭代计算停止准则,且 V 的初始值设为单位矩阵。下面的算法计算矩阵 $A_{m \times n}$ 的奇异值和右奇异向量矩阵 V 。

for all pairs $i < j$

计算矩阵 $A^T A$ 的 (i, j) 子矩阵 $\begin{bmatrix} a & c \\ c & b \end{bmatrix}$:

$$a = \sum_{k=1}^n A_{ki}^2$$

$$b = \sum_{k=1}^n A_{kj}^2$$

$$c = \sum_{k=1}^n A_{ki} * A_{kj}$$

计算使子矩阵 $\begin{bmatrix} a & c \\ c & b \end{bmatrix}$ 对角化的右边 Jacobi 旋转矩阵 $\begin{bmatrix} cs & -sn \\ sn & cs \end{bmatrix}$:

$$\xi = (b - a)/(2c); t = \text{sgn}(\xi)/(|\xi| + \sqrt{1 + \xi^2})$$

$$cs = 1/\sqrt{1 + t^2}; sn = cs * t$$

更新 A 的第 i 列和第 j 列:

for $k = 1 : n$

$$tmp = A_{ki}$$

$$A_{ki} = cs * tmp - sn * A_{kj}$$

$$A_{kj} = sn * tmp + cs * A_{kj}$$

endfor

更新右奇异向量的矩阵 V :

for $k = 1 : n$

$$tmp = V_{ki}$$

$$V_{ki} = cs * tmp - sn * V_{kj}$$

$$V_{kj} = sn * tmp + cs * V_{kj}$$

endfor

endfor

重复以上计算步骤,直到收敛。迭代计算的停止准则是:对所有 $i < j$ 的子矩阵,均有 $|c|/\sqrt{ab} \leq \text{tol}$ 。

上述算法的结果如下:

- (1) 最后得到的矩阵 A 各列的范数给出奇异值的计算结果。
- (2) 矩阵 V 的各列给出与上述奇异值对应的右奇异向量。
- (3) 更新后的矩阵 A 的各列给出与上述奇异值对应的左奇异向量。

注意,虽然我们在第 4 章按照大多数文献的叫法,将 Jacobi 旋转和 Givens 旋转当作完全相同的旋转,但是上述算法中的 Jacobi 旋转确实与 Givens 旋转不同。其不同点在于参数 cs 和 sn 的计算不一样。

6.4 乘积奇异值分解

上一节介绍了一般矩阵的奇异值分解。从本节开始,将依次讨论几种特殊情况下矩阵的奇异值分解,它们分别是乘积奇异值分解、广义奇异值分解和结构奇异值分解。本

节介绍乘积奇异值分解的有关理论和实现算法。

6.4.1 乘积奇异值分解问题

所谓乘积奇异值分解 (product singular value decomposition, PSVD), 顾名思义就是两个矩阵乘积 $B^T C$ 的奇异值分解。考虑矩阵乘积

$$A = B^T C, \quad B \in R^{p \times m}, \quad C \in R^{p \times n}, \quad \text{rank}(B) = \text{rank}(C) = p \quad (6.4.1)$$

从原理上讲, 乘积奇异值分解等价于直接对矩阵的乘积进行普通的奇异值分解。然而, 事先直接计算矩阵的乘积, 再计算矩阵乘积的奇异值分解往往会让小的奇异值产生大的扰动。为了说明这一点, 请看一个例子。

例 6.4.1^[134] 令

$$B^T = \begin{bmatrix} 1 & \xi \\ -1 & \xi \end{bmatrix}, \quad C = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad B^T C = \frac{1}{\sqrt{2}} \begin{bmatrix} 1-\xi & 1+\xi \\ -1-\xi & -1+\xi \end{bmatrix} \quad (6.4.2)$$

显然, C 是一个正交矩阵, 而 B^T 的两列 $[1, -1]^T$ 和 $[\xi, \xi]^T$ 相互正交。矩阵乘积 $B^T C$ 的真实奇异值为 $\sigma_1 = \sqrt{2}$ 和 $\sigma_2 = \sqrt{2}|\xi|$ 。然而, 若 $|\xi|$ 小于截止误差 ε , 式 (6.4.2) 的浮点计算结果为 $B^T C = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$, 其奇异值为 $\sigma_1 = \sqrt{2}$ 和 $\sigma_2 = 0$ 。若 $|\xi| > 1/\varepsilon$, 则浮点运算得到的矩阵乘积 $B^T C = \frac{1}{\sqrt{2}} \begin{bmatrix} -\xi & \xi \\ -\xi & \xi \end{bmatrix}$, 其奇异值为 $\sigma_1 = 0$ 和 $\sigma_2 = \sqrt{2}|\xi|$ 。因此, 矩阵乘积 $B^T C$ 的两个实际的奇异值 $\sigma_1 = \sqrt{2}$ 和 $\sigma_2 = \sqrt{2}|\xi|$ 在经过浮点算法计算后, 最小的奇异值被扰动为 0, 与实际的奇异值相差明显。Laub 等人^[273]指出, 当线性系统接近不可控和不可观测时, 小奇异值的精确计算显得十分重要, 因为如果一个非零的小奇异值被计算为零值, 则会导致错误的结论, 即将一个最小系统判断为非最小系统。

上述例子说明, 直接对两个矩阵的乘积 $B^T C$ 进行奇异值分解在数值上是不可取的。因此, 有必要考虑一个更加困难的问题: 能否使得计算式 (6.4.1) 中 $A = B^T C$ 的奇异值分解尽可能与给定的 B 和 C 具有接近的精度? 这就是所谓的 (矩阵) 乘积奇异值分解问题。

乘积奇异值分解是由 Fernando 与 Hammarling 于 1988 年首先提出来的^[153], 它可以下面的定理来表述。

定理 6.4.1 (乘积奇异值分解)^[153] 令 $B^T \in C^{m \times p}$, $C \in C^{p \times n}$, 则存在酉矩阵 $U \in C^{m \times m}$, $V \in C^{n \times n}$ 和非奇异矩阵 $Q \in C^{p \times p}$ 使得

$$UB^H Q = \begin{bmatrix} I & & \\ & O_B & \\ & & \Sigma_B \end{bmatrix}, \quad Q^{-1} C V^H = \begin{bmatrix} O_C & & \\ & I & \\ & & \Sigma_C \end{bmatrix} \quad (6.4.3)$$

式中

$$\begin{aligned} \Sigma_B &= \text{diag}(s_1, s_2, \dots, s_r), & 1 > s_1 \geq \dots \geq s_r > 0 \\ \Sigma_C &= \text{diag}(t_1, t_2, \dots, t_r), & 1 > t_1 \geq \dots \geq t_r > 0 \end{aligned}$$

且

$$s_i^2 + t_i^2 = 1, \quad i = 1, 2, \dots, r$$

有关本定理的证明, 可参见文献 [153], 但在 6.6 节 (约束奇异值分解) 中, 将介绍一种更简单的证明方法。根据定理 6.4.1, 不难验证

$$UB^H CV^H = \text{diag}(O_C, O_B, \Sigma_B \Sigma_C)$$

因此, 矩阵乘积 $B^H C$ 的奇异值由零奇异值和非零奇异值两部分组成, 其非零奇异值由 $s_i t_i, i = 1, 2, \dots, r$ 给出。

6.4.2 乘积奇异值分解的三角型 Kogbetliantz 算法

定理 6.4.1 给出了乘积奇异值分解的严格形式。值得指出的是, 在文献 [153] 给出这种严格定义之前, 已由 Heath 等人 [208] 于 1986 年提出了矩阵乘积奇异值分解的一种松散形式, 而且 Heath 等人的方法直接给出了一种计算乘积奇异值的实际算法。

给定一个 $p \times m$ 实矩阵 B 和一个 $p \times n$ 实矩阵 C , 并假定它们都具有满秩 $\text{rank}(B) = \text{rank}(C) = p < \min\{m, n\}$ 。即是说, 矩阵乘积 $B^T C$ 有 p 个非零奇异值, 其余 $\min\{m, n\} - p$ 个奇异值全部为零。令

$$B^T = Q_B \begin{bmatrix} R_B \\ O \end{bmatrix}, \quad C = [R_C, O] Q_C \quad (6.4.4)$$

式中, Q_B, Q_C 同为正交矩阵, 且 R_B, R_C 均为上三角矩阵。于是, 矩阵乘积

$$B^T C = Q_B \begin{bmatrix} R_B R_C & O \\ O & O \end{bmatrix} Q_C \quad (6.4.5)$$

上式表明, 矩阵乘积 $B^T C$ 的 p 个非零奇异值与两个上三角矩阵的乘积 $R_B R_C$ 的 p 个非零奇异值相同。

由于矩阵变换为三角矩阵之后, 能够提高奇异值分解的数值性能 [75], 所以计算 $R_B R_C$ 的非零奇异值具有比直接计算 $B^T C$ 的奇异值更好的数值性能, 使得 $R_B R_C$ 的非零奇异值具有与给定的 B 和 C 接近的精度。

Kogbetliantz 算法 [259] 是计算任意正方矩阵奇异值分解的一种最常用的有效方法。这种算法的基本思想是将 $n \times n$ 矩阵 A 的奇异值分解问题转化为一系列 2×2 矩阵的奇异值分解, 也就是用一个迭代的循环序列计算 $n \times n$ 正方矩阵的奇异值分解。以 $n = 4$ 为例, 在序列循环法的每一步, 计算以对角线为中心的 2×2 矩阵的奇异值分解以消去对角线以外的矩阵元素, 循环的顺序如下:

$$\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 4 & 5 \\ 2 & 4 & 0 & 6 \\ 3 & 5 & 6 & 0 \end{bmatrix} \quad (6.4.6)$$

称之为行排序, 当然也有列排序。式中, 标 1 的元素第一个被消去, 标 2 的元素第二个被消去, 如此等等。需要注意的是, 在消去标 2 的元素时, 原先已消去的标 1 元素可能会重

新出现非零值。更一般地,在消去标 $k (\geq 2)$ 的元素时,原消去的标 $(k-1)$ 元素有可能重新出现非零值。当然,对角线上的元素(标 0)不被消去。上述算法可以保证原矩阵 A 收敛为对角矩阵形式。实际中发现,算法至少是二次收敛的。

对于一个三角矩阵 A_1 , Kogbetliantz 算法将它变换为下三角矩阵 A_2 。再次对 A_2 运用 Kogbetliantz 算法,又将下三角矩阵变换成上三角矩阵 A_3 。因此,随着迭代的进行,矩阵 A_k 在上三角和下三角形式之间来回变换。而且,在每一循环中,只有一个矩阵元素被 2×2 奇异值分解消去。以 4×4 上三角矩阵为例,循环顺序如图 6.4.1 所示。

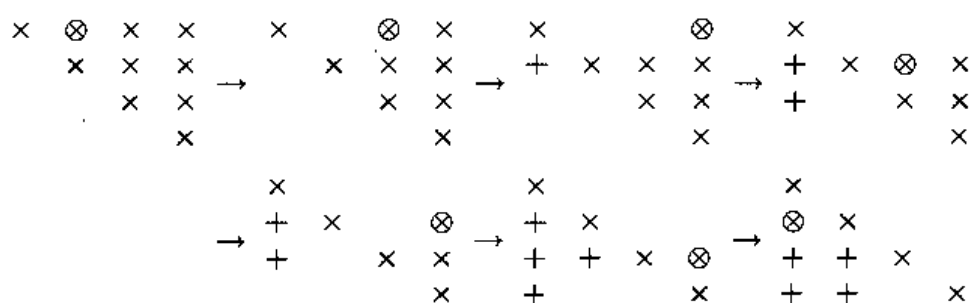


图 6.4.1 利用一次扫零将上三角矩阵变换为下三角矩阵

图 6.4.1 中,原来的非零元素用 \times 标注,新产生的非零元素用 $+$ 表示,而标有 \otimes 的元素则是下一步需要消去的。这样,经过一次扫零,就得到了下三角形式。下一次扫零使用完全相同的顺序,并再一次产生上三角矩阵。

计算乘积奇异值分解的上述方法是 Heath 等人提出的^[208]。通过观察这一方法,可以得出以下结果:

- (1) 这种算法理论上等价于对矩阵 A_k 应用 Kogbetliantz 算法,所以每一步矩阵 A 的非对角线元素理论上都是减小的,并且最终将收敛为零值。
- (2) 每一步只涉及 B_k 的两列和 C_k 的两列被读取或被修正。
- (3) 每一步只计算 A_k 的四个元素。这四个元素的两个总是位于对角线上,它们可以从前一步知道。因此,如果利用对 $\alpha_{ii}^{(k+1)}$ 和 $\alpha_{jj}^{(k+1)}$ 的了解,(1,2) 元素的化简需要计算 4 次向量内积,但是对 (1,3), \dots , (1, n) 各元素的化简却只需要计算 3 次向量内积,而对剩下元素的化简只计算 2 次向量内积。
- (4) 当 $B = C$ 时,上述方法实质上就变成了计算 $A = B^T B$ 奇异值分解的算法(只需要单边的 Jacobi 方法)。这种“单边”奇异值分解方法是 Hestenes 于 1958 年提出的^[216]。因此,Heath 等人采用双边 Kogbetliantz 算法的“双边”奇异值分解方法可以看作是 Hestenes 单边奇异值分解方法的推广。

概而言之,Heath 等人的算法理论上等价于对 $A = B^T C$ 使用 Kogbetliantz 三角型奇异值分解算法。Kogbetliantz 算法使用 QR 分解,将 B_k^T 和 C_k 分别变换为三角矩阵形式,并且在迭代过程中,通过使用平面旋转,能够保持两个矩阵的三角形状。迭代阶段的收敛有可能慢,尤其是当存在多重奇异值,或者奇异值的大小比较接近时,收敛为对角

矩阵的过程更慢。Heath 等人通过数值计算证明, 上述算法具有比直接计算 $B^T C$ 的奇异值分解更好的精确度。

然而, 在某些情况下, Heath 等人的算法会产生与直接计算 $B^T C$ 的奇异值分解相类似的误差。以最小二乘问题中的法方程为例。

例 6.4.2^{[273],[47],[134]} 令

$$B = \begin{bmatrix} 0 & \xi \\ 1 & 1 \end{bmatrix}, \quad C = B, \quad A = B^T C = \begin{bmatrix} 0 & 1 \\ \xi & 1 \end{bmatrix} \begin{bmatrix} 0 & \xi \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 + \xi^2 \end{bmatrix}$$

其真实奇异值为 $\sigma_1 = |\xi|/\sqrt{2}$ 和 $\sigma_2 = \sqrt{2 + 0.5\xi^2}$ 。显然, 若 $|\xi| > \varepsilon$, 但 $|\xi|^2 < \varepsilon$, 则矩阵 A 的浮点近似 $\tilde{A} \approx B^T C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 。由于 B^T 和 C 同为上三角矩阵, 使用 Heath 等人的算法计算的奇异值将与直接计算 \tilde{A} 的奇异值相同, 为 $\hat{\sigma}_1 = 0$ 和 $\hat{\sigma}_2 = \sqrt{2}$ 。显然, 小的奇异值的计算结果 $\hat{\sigma}_1 = 0$ 与真实的奇异值 $\sigma_1 = |\xi|/\sqrt{2}$ 存在较大的误差。

6.4.3 乘积奇异值分解的精确计算

为了克服 Heath 等人的算法的这一缺点, Drmac 于 1998 年提出了乘积奇异值分解的精确计算算法^[134], 其基本思路如下: 任何一个矩阵 A 与正交矩阵相乘, 其奇异值保持不变。因此, 若令

$$B' = TBU, \quad C' = (T^T)^{-1}CV \quad (6.4.7)$$

其中, T 非奇异, U, V 为正交矩阵, 则 $B'^T C' = U^T B^T C V$ 与 $B^T C$ 具有完全相同的奇异值 (包括零奇异值在内), 并且很容易由 $B'^T C'$ 的奇异值分解得到 $B^T C$ 的奇异值分解, 因为

$$B'^T C' = U^T B^T T^T (T^T)^{-1} C V = U^T (B^T C) V$$

给定矩阵 $B \in R^{p \times m}, C \in R^{p \times n}, p \leq \min\{m, n\}$, 并记矩阵 B 的行向量为 $b_i^T, i = 1, 2, \dots, p$ 。Drmac 的乘积奇异值分解算法如下。

算法 6.4.1 (乘积奇异值分解 PSVD(B, C))^[134]

步骤 1 计算 $B_r = \text{diag}(\|b_1^T\|_2, \|b_2^T\|_2, \dots, \|b_p^T\|_2)$, 令 $B_1 = B_r^+ B, C_1 = B_r^+ C$ 。

步骤 2 利用算法 4.7.1 (列旋转的 Householder QR 分解算法) 计算 C_1^T 的 QR 分解, 即

$$C_1^T \Pi = Q \begin{bmatrix} R \\ O_{(n-r) \times p} \end{bmatrix}$$

其中, $R \in R^{r \times p}, \text{rank}(R) = r$; Q 为正交矩阵。

步骤 3 利用标准矩阵乘法计算矩阵 $F = B_1^T \Pi R^T$ 。

步骤 4 计算矩阵 F 的 QR 分解 (最好使用列旋转的 Householder QR 分解算法):

$$F \Pi_F = Q_F \begin{bmatrix} R_F \\ O \end{bmatrix}$$

步骤 5 对转置矩阵 R_F^T 应用奇异值分解的右边 Jacobi 算法 (算法 5.3.1), 计算 R_F 的奇异值分解 $\Sigma = V^T R_F W$ 。

输出 矩阵乘积 $B^T C$ 的奇异值分解结果为

$$\begin{bmatrix} \Sigma \oplus O \\ O \end{bmatrix} = \begin{bmatrix} V^T & \\ & I \end{bmatrix} Q_F^T (B^T C) [Q(W \oplus I_{n-p})]$$

式中, $A \oplus D$ 表示矩阵 A 与 D 的直和。

在上述算法中, 对角矩阵 $D = \text{diag}(d_1, d_2, \dots, d_p)$ 的广义逆矩阵 D^\dagger 仍然为对角矩阵, 其对角元素为 $1/d_i$ ($d_i \neq 0$) 或 0 ($d_i = 0$)。

计算矩阵乘积 $B^T C$ 的奇异值分解的上述算法已被推广到三个矩阵乘积的奇异值分解的精确计算^[136]。令

$$A = B^T S C \quad (6.4.8)$$

式中, $B \in R^{p \times m}$, $S \in R^{p \times q}$, $C \in R^{q \times n}$, $p \leq m$, $q \leq n$ 。

假定正则条件

$$\text{rank}(B) = p, \text{rank}(C) = q, \text{rank}(S) = \rho = \min\{p, q\} \quad (6.4.9)$$

满足这一正则条件的三个矩阵 (B, S, C) 称为正则矩阵三元组 (regular matrix triplet)^[136]。在这种情况下, 矩阵 A 将有 $\min\{m, n\} - \rho = \min\{m, n\} - \min\{p, q\}$ 个确定的零奇异值。现在的问题是, 用尽可能高的相对精度计算其他非零奇异值。

下面是 Drmac 于 2000 年提出的两种算法^[136]。

算法 6.4.2 (三矩阵乘积 $B^T S C$ 的奇异值分解 PSVD(B, S, C))^[136]

输入 $B \in R^{p \times m}$, $S \in R^{p \times q}$, $C \in R^{q \times n}$, $p \leq m$, $q \leq n$

步骤 1 计算 $B_\tau = \text{diag}(\|b_1^\tau\|_2, \|b_2^\tau\|_2, \dots, \|b_p^\tau\|_2)$, $C_\tau = \text{diag}(\|c_1^\tau\|_2, \|c_2^\tau\|_2, \dots, \|c_q^\tau\|_2)$, 其中, b_i^τ ($i = 1, 2, \dots, p$) 和 c_j^τ ($j = 1, 2, \dots, q$) 分别是矩阵 B 和 C 的行向量。然后, 令 $B_1 = B_\tau^\dagger B$, $C_1 = C_\tau^\dagger C$, $S_1 = B_\tau S C_\tau$ 。

步骤 2 利用行和列旋转计算矩阵 S_1 的 LU 分解:

$$\Pi_1 S_1 \Pi_2 = LU$$

式中

$$L \in R^{p \times p}, U \in R^{p \times q}, \rho = \text{rank}(L) = \text{rank}(U), L_{ii} = 1, 1 \leq i \leq \rho$$

步骤 3 利用标准的矩阵乘法运算计算

$$M = L^T \Pi_1 B_1, \quad N = U \Pi_2^T C_1$$

应用算法 6.4.1 直接得到 $M^T N$ 的奇异值分解。

输出 三矩阵乘积 $B^T S C$ 的奇异值分解为

$$\begin{bmatrix} \Sigma \oplus O \\ O \end{bmatrix} = \begin{bmatrix} V^T & \\ & I \end{bmatrix} Q_F^T (B^T S C) (Q(W \oplus I_{n-p}))$$

式中, Q, Q_F, V 和 W 为在步骤 3 中使用算法 6.4.1 得到的结果。

算法 6.4.3 (三矩阵乘积 $B^T S^{-1} C$ 的奇异值分解 $\text{PSVD}(B, S^{-1}, C)$)^[136]

输入 $B \in R^{p \times m}$, $S \in R^{p \times p}$, $C \in R^{p \times n}$, $\text{rank}(S) = p$

步骤 1 计算

$$B_\tau = \text{diag}(\|b_1^\tau\|_2, \|b_2^\tau\|_2, \dots, \|b_p^\tau\|_2)$$

$$C_\tau = \text{diag}(\|c_1^\tau\|_2, \|c_2^\tau\|_2, \dots, \|c_p^\tau\|_2)$$

其中, $b_i^\tau (i = 1, 2, \dots, p)$ 和 $c_j^\tau (j = 1, 2, \dots, q)$ 分别是矩阵 B 和 C 的行向量。然后, 令 $B_1 = B_\tau^{-1} B$, $C_1 = C_\tau^{-1} C$, $S_1 = C_\tau^{-1} S B_\tau^{-1}$ 。

步骤 2 利用行和列旋转计算矩阵 S_1 的 LU 分解:

$$\Pi_1 S_1 \Pi_2 = LU, \quad L_{ii} = 1, \quad 1 \leq i \leq p$$

步骤 3 利用标准的矩阵乘法运算计算

$$M = U^{-T} \Pi_2 B_1, \quad N = L^{-1} \Pi_1^T C_1$$

应用算法 6.4.1 直接得到 $M^T N$ 的奇异值分解。

输出 三矩阵乘积 $B^T S^{-1} C$ 的奇异值分解为

$$\begin{bmatrix} \Sigma \oplus O \\ O \end{bmatrix} = \begin{bmatrix} V^T & \\ & I \end{bmatrix} Q_F^T (B^T S^{-1} C) (Q(W \oplus I_{n-p}))$$

式中, Q, Q_F, V 和 W 为在步骤 3 中使用算法 6.4.1 得到的结果。

6.5 广义奇异值分解

6.2 节介绍了矩阵 A 的奇异值分解。矩阵的奇异值分解已有多种推广形式, 6.4 节的矩阵乘积的奇异值分解只是其中之一。从这节起, 我们将讨论其他形式的奇异值分解, 它们是矩阵束 (A, B) 的奇异值分解 (称为广义奇异值分解), 矩阵三元组 (A, B, C) 的奇异值分解 (约束奇异值分解), 结构奇异值分解和 Takagi 奇异值分解。由于奇异值分解有多种推广, 所以有些文献把矩阵 A 的奇异值分解称为普通奇异值分解 (ordinary SVD, OSVD)。本节先讨论广义奇异值分解 (generalized SVD, GSVD), 其他几种推广的奇异值分解则留待 6.6 节再介绍。

6.5.1 对称正定问题

在广义特征值问题 $Ax = \lambda Bx$ 中, 满足广义特征值 - 特征向量方程

$$Ax = \lambda Bx$$

的标量 λ 和非零向量 x 分别称为矩阵束 (A, B) 的广义特征值和广义特征向量。求解广义特征值问题 $Ax = \lambda Bx$ 的方法之一是下面的全等变换:

$$A - \lambda B \quad (X^T A X) - \lambda (X^T B X)$$

\Longleftrightarrow

对称 - 正定

对称 - 正定

并且希望寻求一种稳定的有效算法, 它计算的 X 能使得 $X^T A X$ 和 $X^T B X$ 二者都是“标准型”的。最简单的标准型就是对角矩阵。

定理 6.5.1 假定 A 和 B 为 $n \times n$ 对称矩阵, 并定义

$$C(\mu) = \mu A + (1 - \mu)B, \quad \mu \in R \quad (6.5.1)$$

若存在 $\mu \in [0, 1]$ 使得 $C(\mu)$ 是非负定的, 并且

$$\text{Null}(C(\mu)) = \text{Null}(A) \cap \text{Null}(B) \quad (6.5.2)$$

则存在一个非奇异矩阵 Q 使得 $Q^T A Q$ 和 $Q^T B Q$ 二者都是对角形式的。

证明 选择 $\mu \in [0, 1]$ 使 $C(\mu)$ 非负定, 并具有性质 (6.5.2) 式。设

$$Q_0^T C(\mu) Q_0 = \begin{bmatrix} D & O \\ O & O \end{bmatrix}, \quad D = \text{diag}(d_1, d_2, \dots, d_k), \quad d_i > 0$$

是 $C(\mu)$ 的 Schur 分解, 并定义 $Q_1 = Q_0 \text{diag}(D^{-1/2}, I_{n-k})$ 。如果 $A_1 = Q_1^T A Q_1$, $B_1 = Q_1^T B Q_1$ 和 $C_1 = Q_1^T C(\mu) Q_1$, 则

$$C_1 = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} = \mu A_1 + (1 - \mu)B_1$$

由于 $\text{Span}\{e_{k+1}, e_{k+2}, \dots, e_n\} = \text{Null}(C_1) = \text{Null}(A_1) \cap \text{Null}(B_1)$, 故知 A_1 和 B_1 具有以下结构:

$$A_1 = \begin{bmatrix} A_{11} & O \\ O & O \end{bmatrix}, \quad B_1 = \begin{bmatrix} B_{11} & O \\ O & O \end{bmatrix}$$

其中, A_{11} 和 B_{11} 均是 $k \times k$ 方阵, 且 $I_k = \mu A_{11} + (1 - \mu)B_{11}$ 。

假定 $\mu \neq 0$ 。于是, 若 $Z^T B_{11} Z = \text{diag}(b_1, b_2, \dots, b_k)$ 是 B_{11} 的 Schur 分解, 则可通过令 $Q = Q_1 \text{diag}(Z, I_{n-k})$, 得到

$$Q^T B Q = \text{diag}(b_1, \dots, b_k, 0, \dots, 0) \equiv D_B$$

和

$$Q^T A Q = \frac{1}{\mu} Q^T [C(\mu) - (1 - \mu)B] Q = \frac{1}{\mu} \left(\begin{bmatrix} I_k & O \\ O & O \end{bmatrix} - (1 - \mu)D_B \right) \equiv D_A$$

另外一方面, 若 $\mu = 0$, 则令 $Z^T A_{11} Z = \text{diag}(a_1, a_2, \dots, a_k)$ 是 A_{11} 的 Schur 分解, 并令 $Q = Q_1 \text{diag}(Z, I_{n-k})$ 。容易验证, 在这种情况下, 矩阵 $Q^T A Q$ 和 $Q^T B Q$ 二者也都是对角的。 ■

对于对称 - 正定的矩阵束, 上述定理给出下面的结果。

推论 6.5.1 若 $A - \lambda B$ 是对称 - 正定的, 则存在一个非奇异矩阵 $Q = [q_1, q_2, \dots, q_n]$ 使得

$$Q^T A Q = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \quad Q^T B Q = \text{diag}(\beta_1, \beta_2, \dots, \beta_n)$$

且 $Aq_i = \lambda_i Bq_i, i = 1, 2, \dots, n$, 其中, $\lambda_i = \alpha_i / \beta_i$ 为广义特征值。

证明 在定理 6.5.1 中令 $\mu = 0$ 后立即知, 对称 - 正定矩阵束可以同时对角化。推论的其他部分很容易建立。 ■

6.5.2 广义奇异值分解

上一小节的对称 - 正定问题实际上又可以写作 $A^T A - \lambda B^T B$ 的形式, 其中, $A \in R^{m \times n}$, $B \in R^{p \times n}$ 。注意, 由定理 6.5.1 知, 存在一个非奇异矩阵 $Q \in R^{n \times n}$ 使得 $Q^T(A^T A)Q$ 和 $Q^T(B^T B)Q$ 二者都是对角的。因此, 无需事先计算 $A^T A$ 和 $B^T B$, 就有可能直接由 A 和 B 求出广义特征值 λ 。这就是所谓的广义奇异值分解 (GSVD) 方法, 它是 Van Loan 于 1976 年最早提出来的 [462]。

定理 6.5.2 (广义奇异值分解 1) [462] 若 $A \in C^{m \times n}, m \geq n$ 和 $B \in C^{p \times n}$, 则存在酉矩阵 $U \in C^{m \times m}$ 和 $V \in C^{p \times p}$ 以及非奇异矩阵 $Q \in C^{n \times n}$, 使得

$$U A Q = \begin{bmatrix} k & n-k \\ \Sigma_A & O \end{bmatrix}, \quad \Sigma_A = \begin{bmatrix} I_r & & \\ & S_A & \\ & & O_A \end{bmatrix} \quad (6.5.3)$$

$$V B Q = \begin{bmatrix} k & n-k \\ \Sigma_B & O \end{bmatrix}, \quad \Sigma_B = \begin{bmatrix} O_B & & \\ & S_B & \\ & & I_{k-r-s} \end{bmatrix} \quad (6.5.4)$$

式中

$$S_A = \text{diag}(\alpha_{r+1}, \alpha_{r+2}, \dots, \alpha_{r+s}), \quad S_B = \text{diag}(\beta_{r+1}, \beta_{r+2}, \dots, \beta_{r+s}) \quad (6.5.5)$$

$$\left. \begin{aligned} 1 &> \alpha_{r+1} \geq \dots \geq \alpha_{r+s} > 0 \\ 0 &< \beta_{r+1} \leq \dots \leq \beta_{r+s} < 1 \\ \alpha_i^2 + \beta_i^2 &= 1, \quad i = r+1, r+2, \dots, r+s \end{aligned} \right\} \quad (6.5.6)$$

整数 k, r 和 s 分别为

$$k = \text{rank} \begin{bmatrix} A \\ B \end{bmatrix}, \quad r = \text{rank} \begin{bmatrix} A \\ B \end{bmatrix} - \text{rank}(B)$$

和

$$s = \text{rank}(A) + \text{rank}(B) - \text{rank} \begin{bmatrix} A \\ B \end{bmatrix}$$

证明 本定理有多种证明方法, 可参见 Van Loan [462], Paige 与 Saunders [424], Golub 与 Van Loan [184] 和 Zha [499] 的证明。这里给出的是 Zha 的构造性证明, 它由四步组成, 每一步都是下列形式的变换:

$$A^{(k+1)} = U^{(k)} A^{(k)} Q^{(k)}, \quad B^{(k+1)} = V^{(k)} B^{(k)} Q^{(k)}$$

其中, $U^{(k)}$ 和 $V^{(k)}$ 为酉矩阵, 且 $Q^{(k)}$ 非奇异. 在每一步, 只要规定 $U^{(k)}, V^{(k)}, Q^{(k)}$ 以及变换后得到的矩阵 $A^{(k+1)}, B^{(k+1)}$. 令 $A^{(1)} = A$ 和 $B^{(1)} = B$.

步骤 1 令矩阵 B 的奇异值分解为 $U_1 B V_1 = \text{diag}(0, \Sigma_B^{(1)})$, 其中, $\Sigma_B^{(1)} = \text{diag}(s_1, s_2, \dots, s_t)$, 并且 $s_1 \geq s_2 \geq \dots \geq s_t > 0$. 令

$$U^{(1)} = I, \quad V^{(1)} = U_1, \quad Q^{(1)} = V_1 \text{diag}(I, \Sigma_B^{-1})$$

则

$$A^{(2)} = [A_1^{(2)} \quad A_2^{(2)}], \quad B^{(2)} = \begin{bmatrix} O & O \\ O & I_t \end{bmatrix}$$

步骤 2 令矩阵 $A_1^{(2)}$ 的奇异值分解为 $U_2 A_1^{(2)} V_2 = \text{diag}(\Sigma_A^{(2)}, O)$, 其中, $\Sigma_A^{(2)} = \text{diag}(t_1, t_2, \dots, t_r)$, 且 $t_1 \geq t_2 \geq \dots \geq t_r > 0$. 设

$$U^{(2)} = U_2, \quad V^{(2)} = I, \quad Q^{(2)} = \text{diag}(V_2, I) \text{diag}((\Sigma_A^{(2)})^{-1}, I)$$

则

$$A^{(3)} = \begin{bmatrix} I_r & O & A_{13}^{(3)} \\ O & O & A_{23}^{(3)} \end{bmatrix} \begin{matrix} r & n-r-t & t \\ m-r & & \end{matrix}, \quad B^{(3)} = B^{(2)}$$

步骤 3 令 $A_{23}^{(3)}$ 的奇异值分解为 $U_3 A_{23}^{(3)} V_3 = \text{diag}(\Sigma_A^{(3)}, O)$, 其中, $\Sigma_A^{(3)} = \text{diag}(w_1, w_2, \dots, w_s)$, 且 $w_1 \geq w_2 \geq \dots \geq w_s > 0$. 设 $\alpha_i = w_i(1+w_i^2)^{-1/2}$, $\beta_i = (1+w_i^2)^{-1/2}$, $i = r+1, r+2, \dots, r+s$ 和 $S_A = \text{diag}(\alpha_{r+1}, \alpha_2, \dots, \alpha_{r+s})$, $S_B = \text{diag}(\beta_{r+1}, \beta_{r+2}, \dots, \beta_{r+s})$. 容易验证, α_i 和 β_i ($i = r+1, r+2, \dots, r+s$) 满足式 (6.5.6). 令

$$U^{(3)} = \text{diag}(I, U_3)$$

$$V^{(3)} = \text{diag}(I, V_3^H)$$

$$Q^{(3)} = \begin{bmatrix} I & -A_{13}^{(3)} \\ O & I \end{bmatrix} \text{diag}(I, V_3) \text{diag}(I, S_B, I)$$

则

$$A^{(4)} = \begin{bmatrix} I_r & O & O & O \\ O & O & S_A & O \\ O & O & O & O \end{bmatrix} \begin{matrix} r & n-r-t & s & t-s \\ s & & & \\ m-r-s & & & \end{matrix}$$

和

$$B^{(4)} = \begin{bmatrix} O & O & O \\ O & S_B & O \\ O & O & I_{k-r-s} \end{bmatrix} \begin{matrix} n-k+r & s & t-s \\ s & & \\ k-r-s & & \end{matrix}$$

步骤 4 经过适当的矩阵置换 P_1 和 P_2 , 并令 $k = t + r$ 后, 我们得到

$$A^{(5)} = A^{(4)} P_1 = \begin{bmatrix} I_r & & & \\ & S_A & & \\ & & O_A & \\ & & & \vdots & \\ & & & & O \end{bmatrix}$$

以及

$$B^{(5)} = P_2 B^{(4)} P_1 = \begin{bmatrix} O_B & & & \\ & S_B & & \\ & & I_{k-r-s} & \\ & & & \vdots & \\ & & & & O \end{bmatrix}$$

于是, 就得到了所需要的伪对角形式。容易验证:

$$\text{rank}(A) = r + s, \quad \text{rank}(B) = k - r, \quad \text{rank} \begin{bmatrix} A \\ B \end{bmatrix} = k$$

于是, 本定理得证。■

根据文献 [462], 式 (6.5.3) 的对角矩阵 Σ_A 和式 (6.5.4) 的对角矩阵 Σ_B 的对角线上的元素组成广义奇异值对 (α_i, β_i) 。由 Σ_A 和 Σ_B 的形式, 前 k 个广义奇异值对分为三种情况:

$$\alpha_i = 1, \quad \beta_i = 0, \quad i = 1, 2, \dots, r$$

$$\alpha_i, \beta_i \quad (S_A \text{ 和 } S_B \text{ 的元素}), \quad i = r+1, r+2, \dots, r+s$$

$$\alpha_i = 0, \quad \beta_i = 1, \quad i = r+s+1, r+s+2, \dots, k$$

这 k 个奇异值对 (α_i, β_i) 统称矩阵束 (A, B) 的非平凡广义奇异值对; 而 α_i/β_i ($i = 1, 2, \dots, k$) 称为矩阵束 (A, B) 的非平凡广义奇异值 (包括无穷大, 有限值和零)。反之, 对应于式 (6.5.3) 和式 (6.5.4) 中零列向量的另外 $n-k$ 对广义奇异值则称为矩阵束 (A, B) 的平凡广义奇异值对。

定理 6.5.1 限制矩阵 A 的列数不得大于行数。当矩阵 A 的维数不满足这一限制时, 定理 6.5.1 便不能适用。Paige 与 Saunders^[349] 推广了定理 6.5.1, 提出了具有相同列数的任意矩阵束 (A, B) 的广义奇异值分解。

定理 6.5.3 (广义奇异值分解 2)^[349] 假定矩阵 $A \in C^{m \times n}$ 和 $B \in C^{p \times n}$, 则对于分块矩阵

$$K = \begin{bmatrix} A \\ B \end{bmatrix}, \quad t = \text{rank}(K)$$

存在酉矩阵

$$U \in C^{m \times m}, \quad V \in C^{p \times p}, \quad W \in C^{t \times t}, \quad Q \in C^{n \times n}$$

使得

$$\begin{aligned} U^H A Q &= \Sigma_A \left[\underbrace{W^H R}_t, \underbrace{O}_{n-t} \right] \\ V^H B Q &= \Sigma_B \left[\underbrace{W^H R}_t, \underbrace{O}_{n-t} \right] \end{aligned}$$

式中

$$\Sigma_A = \begin{bmatrix} I_A & & \\ & D_A & \\ & & O_A \end{bmatrix}, \quad \Sigma_B = \begin{bmatrix} I_B & & \\ & D_B & \\ & & O_B \end{bmatrix} \quad (6.5.7)$$

并且 $R \in C^{t \times t}$ 非奇异, 其奇异值等于矩阵 K 的非零奇异值. 矩阵 I_A 为 $r \times r$ 单位矩阵, I_B 为 $(t-r-s) \times (t-r-s)$ 单位矩阵, 其中, r 和 s 的值与所给数据有关, 且 O_A 和 O_B 分别为 $(m-r-s) \times (t-r-s)$ 维和 $(p-t+r) \times r$ 维零矩阵 (这两个零矩阵有可能没有任何行或任何列), 而

$$D_A = \text{diag}(\alpha_{r+1}, \alpha_{r+2}, \dots, \alpha_{r+s}), \quad D_B = \text{diag}(\beta_{r+1}, \beta_{r+2}, \dots, \beta_{r+s})$$

满足

$$1 > \alpha_{r+1} \geq \alpha_{r+2} \geq \dots \geq \alpha_{r+s} > 0, \quad 0 < \beta_{r+1} \leq \beta_{r+2} \leq \dots \leq \beta_{r+s} < 1$$

和

$$\alpha_i^2 + \beta_i^2 = 1, \quad i = r+1, r+2, \dots, r+s$$

证明 参见文献 [231].

求解广义特征值问题 $A^T A - \lambda B^T B = 0$ 等价于求解 $\det(A^T A - \lambda B^T B) = 0$. 令 $m \times n$ 矩阵 A 和 $p \times n$ 矩阵 B 的奇异值对是 (α, β) , 并代入 $\lambda = \alpha/\beta$, 则广义特征值问题又可进一步等价于求解 $\det(\beta^2 A^H A - \alpha^2 B^H B) = 0$. 当矩阵 B 为正方矩阵, 并且非奇异时, 则有

$$\det(\beta^2 A^H A - \alpha^2 B^H B) = 0 \iff \det[(AB^{-1})^H AB^{-1} - \alpha^2/\beta^2] = 0$$

这说明, 当 B 为非奇异的正方矩阵时, 矩阵束 (A, B) 的广义奇异值分解等价于矩阵 AB^{-1} 的奇异值分解.

下面是有关广义奇异值分解的几点注释.

注释 1 由 (A, B) 的广义奇异值分解与 AB^{-1} 的奇异值分解之间的等价性显见, 若矩阵 B 为单位矩阵 ($B = I$), 则广义奇异值分解简化为普通的奇异值分解. 这一观察结果也可从广义奇异值的定义直接得出. 这是因为, 单位矩阵的奇异值全部等于 1, 从而矩阵束 (A, I) 的广义奇异值与 A 的奇异值等价.

注释 2 由于 AB^{-1} 具有类似于商的形式, 以及广义奇异值本身就是矩阵 A 和 B 的奇异值之商, 所以广义奇异值分解有时也被称作商奇异值分解 (quotient singular value decomposition, QSVD).

注释 3 如果矩阵 B 不是正方的, 或者 B 是奇异的正方矩阵, 则 AB^\dagger (其中, B^\dagger 是 B 的 Moore-Penrose 广义逆) 的奇异值不一定对应为矩阵束 (A, B) 的广义奇异值. 更严格地, 有以下结论.

定理 6.5.4 [499] 定义

$$B_A^\dagger = Q \begin{bmatrix} O_B^H & & \\ & S_B^{-1} & \\ & & I \end{bmatrix} V$$

若 $\text{rank}[A^H, B^H]^H = n$, 则 B_A^\dagger 是唯一定义的, 并且 AB_A^\dagger 的奇异值包含了矩阵束 (A, B) 的全部有限大的广义奇异值。

证明^[499] 由于 $\text{rank}[A^H, B^H]^H = n$, 定理 6.5.2 中任何两组变换都满足下列关系:

$$Q_1 = Q_2 \text{diag}(U_{11}, U_{22}, V_{23})$$

$$U_1^H = U_2^H \text{diag}(U_{11}, U_{22}, V_{23})$$

$$V_1^H = V_2^H \text{diag}(U_{11}, U_{22}, V_{23})$$

因此

$$\begin{aligned} Q_1 \begin{bmatrix} O_B^H & & \\ & S_B^{-1} & \\ & & I \end{bmatrix} V_1 \\ &= Q_2 \begin{bmatrix} U_{11} & & \\ & U_{22} & \\ & & V_{33} \end{bmatrix} \begin{bmatrix} O_B^H & & \\ & S_B^{-1} & \\ & & I \end{bmatrix} \begin{bmatrix} V_{11} & & \\ & U_{22} & \\ & & V_{33} \end{bmatrix} V_2 \\ &= Q_2 \begin{bmatrix} O_B^H & & \\ & S_B^{-1} & \\ & & I \end{bmatrix} V_1 \end{aligned}$$

这就证明了, B_A^\dagger 不存在病态问题。进一步地, 我们观察

$$UAB_A^\dagger V^H = \text{diag}(O, S_A, S_B^{-1}, O)$$

只有矩阵束 (A, B) 的无穷大的广义奇异值变成了矩阵 AB_A^\dagger 的零奇异值, 而其他的所有广义奇异值保留在 AB_A^\dagger 的奇异值内。■

下面进一步分析 B_A^\dagger 的性质。容易验证, B_A^\dagger 满足下列等式:

$$BB_A^\dagger B = B \quad (6.5.8)$$

$$B_A^\dagger BB_A^\dagger = B_A^\dagger \quad (6.5.9)$$

$$(BB_A^\dagger)^H = BB_A^\dagger \quad (6.5.10)$$

注意, 矩阵 B_A^\dagger 是只满足 Moore-Penrose 广义逆矩阵条件 (1), (2) 和 (3), 而不满足条件 (4) 的正规化广义逆矩阵。

下面的定理^[499]表明, B_A^\dagger 是一种约束最小化问题的唯一解。

定理 6.5.5 若 $[A^H, B^H]^H$ 是满列秩的, 则 B_A^\dagger 是下列约束极小化问题的唯一解:

$$\min_{X \in C^{n \times q}} \|AX\|_F \quad (6.5.11)$$

服从约束条件

$$BXB = B \quad (6.5.12)$$

$$XBX = X \quad (6.5.13)$$

$$(BX)^H = BX \quad (6.5.14)$$

$\|AX\|_F$ 的极小化值为 $\sqrt{\sum_{i=r+1}^{r+s} (\alpha_i/\beta_i)^2}$ 。

证明^[499] 令 B 具有奇异值分解

$$VBQ = [\Sigma_B, O]$$

其中, Σ_B 有 k 列。由于 $\text{rank}[A^H, B^H]^H = n$, 所以 $k = n$, 并且 $B = V^H \Sigma_B Q^{-1}$ 。作矩阵分块 $Q^{-1}XV^H = (X_{ij})_{i,j=1}^3$, 并使之适应 Σ_A 和 Σ_B 的分块形式, 则可证明, 为了满足约束条件式 (6.5.12)~式 (6.5.14), X 应具有以下形式:

$$X = Q \begin{bmatrix} O & X_{12} & X_{13} \\ O & S_B^{-1} & O \\ O & O & I_{n-r-s} \end{bmatrix} V$$

这是因为

$$\begin{aligned} \|AX\|_F^2 &= \|UAQQ^{-1}XV^H\|_F^2 \\ &= \left\| \begin{bmatrix} I_r & & \\ & S_A & \\ & & O \end{bmatrix} \begin{bmatrix} O & X_{12} & X_{13} \\ O & S_B^{-1} & O \\ O & O & I_{n-r-s} \end{bmatrix} \right\|_F^2 \\ &= \|[X_{12} \ X_{13}]\|_F^2 + \|S_A S_B^{-1}\|_F^2 \\ &\geq \|S_A S_B^{-1}\|_F^2 \\ &= \sum_{i=r+1}^{r+s} \left(\frac{\alpha_i}{\beta_i} \right)^2 \end{aligned}$$

等式成立, 当且仅当 $X_{12} = O$ 和 $X_{13} = O$, 即 $X = B_A^\dagger$ 。 ■

6.5.3 广义奇异值分解的实际算法

如果 A 或 B 相对于方程求解是病态的, 那么计算 AB^{-1} 通常会导致非常大的数值误差, 所以对 AB^{-1} 本身进行奇异值分解一般并不值得推荐采用。一个自然会问的问题是, 能否绕开计算 AB^{-1} 这一步, 而直接得到 $C = AB^{-1}$ 的奇异值分解? 这是完全可能的, 因为 $C = AB^{-1}$ 的奇异值分解实质上就是两个矩阵乘积的奇异值分解。

Paige^[351] 根据 $C = AB^{-1}$ 的奇异值分解与矩阵乘积的奇异值分解形式上的一致, 提出了一种实际的广义奇异值分解算法。这种算法的关键是如何避免矩阵求逆 B^{-1} 以及如何适用于矩阵 B 奇异的一般情况。

先讨论矩阵 B 非奇异的情况。令 A_{ij} 和 B_{ij} 均代表 2×2 矩阵, 它们的元素分别位于 A 的第 i, j 行和 B 的第 i, j 列。如果选择酉矩阵 U 和 V 使得

$$U^H A_{ij} B_{ij}^{-1} V = S \quad (6.5.15)$$

是对角矩阵, 则

$$U^H A_{ij} = S V^H B_{ij} \quad (6.5.16)$$

结果是, $U^H A_{ij}$ 的第 1 行与 $V^H B_{ij}$ 的第 1 行平行, $U^H A_{ij}$ 的第 2 行与 $V^H B_{ij}$ 的第 2 行平行。因此, 如果 Q 是使得 $V^H B_{ij} Q$ 为下三角矩阵的正交阵, 即

$$(V^H B_{ij})Q = \begin{bmatrix} \times & \otimes \\ \times & \times \end{bmatrix} = \begin{bmatrix} \times & \\ \times & \times \end{bmatrix} \quad (6.5.17)$$

则 $U^H A_{ij} Q$ 也是下三角矩阵。对于 $n \times n$ 上三角矩阵 $C = AB^{-1}$, 可以执行 $n(n-1)/2$ 次 2×2 Kogbetliantz 算法, 使矩阵 A, B 和 C 在上三角和下三角形式之间来回变换, 最后收敛为对角矩阵形式。

广义奇异值分解也可等价叙述为以下定理 [184]。

定理 6.5.6 若 $A \in C^{m_1 \times n} (m_1 \geq n)$, $B \in C^{m_2 \times n} (m_2 \geq n)$, 则存在一非奇异矩阵 $X \in C^{n \times n}$ 使得

$$X^H (A^H A) X = D_A = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \quad \alpha_k \geq 0$$

和

$$X^H (B^H B) X = D_B = \text{diag}(\beta_1, \beta_2, \dots, \beta_n), \quad \beta_k \geq 0$$

式中, $\sigma_k = \sqrt{\alpha_k / \beta_k}$ 称为矩阵束 (A, B) 的广义奇异值, 且 X 的列 x_k 称为与 σ_k 对应的广义奇异向量。

定理 6.5.6 给出了计算矩阵束 (A, B) 的广义奇异值分解的多种算法。特别地, 我们对寻求使 D_B 为单位矩阵的广义奇异向量矩阵 X 更加感兴趣, 因为在这一情况下, 广义奇异值 σ_k 由 $\sqrt{\alpha_k}$ 直接给出。下面就是这样的两种实际算法。

算法 6.5.1 (GSVD 算法 1) [463]

步骤 1 计算矩阵的内积 $S_1 = A^H A$ 和 $S_2 = B^H B$ 。

步骤 2 计算 S_2 的特征值分解 $U_2^H S_2 U_2 = D = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_n)$ 。

步骤 3 计算 $Y = U_2 D^{-1/2}$ 和 $C = Y^H S_1 Y$ 。

步骤 4 计算 C 的特征值分解 $Q^H C Q = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$, 其中, $Q^H Q = I$ 。

步骤 5 广义奇异向量矩阵为 $X = YQ$, 且广义奇异值为 $\sqrt{\alpha_k}$, $k = 1, 2, \dots, n$ 。

证明 直接计算给出

$$X^H (A^H A) X = Q^H Y^H S_1 Y Q = Q^H C Q = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$$

由步骤 2 的结果, 我们有

$$Y^H (B^H B) Y = (U_2 D^{-1/2})^H S_2 (U_2 D^{-1/2}) = I$$

由于 $Q^H Q = I$, 所以上式分别左乘 Q^H 和右乘 Q 后, 则有

$$X^H (B^H B) X = Q^H Y^H (B^H B) Y Q = I \quad \blacksquare$$

算法 6.5.2 (GSVD 算法 2) [463]

步骤 1 计算 B 的奇异值分解 $U_2^H B V_2 = D = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_n)$ 。

步骤 2 计算 $Y = V_2 D^{-1} V_2 = \text{diag}(1/\gamma_1, 1/\gamma_2, \dots, 1/\gamma_n)$ 。

步骤 3 计算 $C = AY$ 。

步骤 4 计算矩阵 C 的奇异值分解 $U_1^H C V_1 = D_A = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$ 。

步骤 5 $X = YV_1$ 为广义奇异向量矩阵, 而 $\alpha_k, k = 1, 2, \dots, n$ 直接是矩阵束 (A, B) 的广义奇异值。

证明 根据算法各步骤的计算结果, 有

$$X^H(A^H A X) = V_1^H (AY)^H (AY) V_1 = V_1^H (C^H C) V_1 = \text{diag}(\alpha_1^2, \alpha_2^2, \dots, \alpha_n^2)$$

对 $Y^H(B^H B)Y = I$ 左乘 V_1^H 和右乘 V_1 得到

$$X^H(B^H B)X = V_1^H I V_1 = I$$

这就证明了 $X^H(A^H A)X$ 和 $X^H(B^H B)X$ 分别具有所希望的形式。■

算法 6.5.1 与算法 6.5.2 的主要区别在于: 前者需要计算矩阵乘积 $A^H A$ 和 $B^H B$, 而后者则完全避免了这一计算。正如前面已说明的那样, 在计算两个矩阵乘积时会发生信息的丢失, 并会使条件数变坏。因此, 算法 6.5.2 具有比算法 6.5.1 更好的数值性能。但是, 由于需要矩阵求逆或矩阵乘积的计算, 算法 6.5.1 和算法 6.5.2 的性能或多或少都会遭到损害。

一种可以避免任何矩阵求逆或矩阵内积运算的广义奇异值分解算法由 Speiser 与 Van Loan [424] 提出 (也见文献 [463])。

算法 6.5.3 (GSVD 算法 3)

步骤 1 计算 QR 分解

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} R$$

其中, Q_1 和 Q_2 分别与 A 和 B 具有相同的维数, 且 $R \in C^{n \times n}$ 为上三角矩阵。假定 R 非奇异, 即 $\text{Null}(A) \cap \text{Null}(B) = \{0\}$ 。

步骤 2 计算 CS 分解

$$\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} U_1 & O \\ O & U_2 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} V$$

其中, U_1, U_2 和 V 为酉矩阵, $C = \text{diag}(\cos(\theta_k))$, $S = \text{diag}(\sin(\theta_k))$, 且 $0 \leq \theta_1 \leq \dots \leq \theta_n \leq \pi/2$ 。由此可知, 若 $X = R^{-1}V$, 则 $X^H(A^H A - \mu^2 B^H B)X = C^H C - \lambda S^H S$, 因此, 广义奇异值由 $\mu_k = \cot(\theta_k)$ 给出。

步骤 3 利用 $c_d > \epsilon + c_n \geq c_{d+1} \geq \dots \geq c_n \geq 0$ ($\epsilon > 0$ 为小的扰动), 其中, $c_k = \cos(\theta_k)$ 。

步骤 4 计算乘积 $ZT = R^H V$ 的 QR 分解, 其中, $Z = [z_1, z_2, \dots, z_n]$ 为酉矩阵, $T \in C^{n \times n}$ 为上三角矩阵。由于

$$X = R^{-1}V = (V^H R)^{-1} = (R^H V)^{-H} = (ZT)^{-H} = ZT^{-H}$$

且 T^{-H} 为下三角矩阵, 故有 $\text{Span}\{z_{d+1}, z_{d+2}, \dots, z_n\} = \text{Span}\{x_{d+1}, x_{d+2}, \dots, x_n\}$ 。

定理 6.5.7 由算法 6.5.3 产生的向量 $\{z_{d+1}, z_{d+2}, \dots, z_n\}$ 准确张成广义奇异值分解问题 $\tilde{A}^H \tilde{A} - \lambda \tilde{B}^H \tilde{B}$ 的最小奇异值子空间, 其中

$$\left\| \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix} - \begin{bmatrix} A \\ B \end{bmatrix} \right\|_F \leq \epsilon \|R\|_F$$

其中, R 由步骤 4 的 QR 分解给出。

证明 定义 $\tilde{C} = \text{diag}(\cos \tilde{\theta}_k)$, $\tilde{S} = \text{diag}(\sin \tilde{\theta}_k)$, 其中, $\tilde{\theta}_k = \theta_k$ 。令 $\tilde{A} = U_1 \tilde{C} V^H R$, $B = U_2 \tilde{S} V^H R$ 。由于

$$\tilde{A} - A = U_1 (\tilde{C} - C) V^H R$$

和

$$\tilde{B} - B = U_2 (\tilde{S} - S) V^H R$$

所以

$$\left\| \begin{bmatrix} \tilde{A} \\ \tilde{B} \end{bmatrix} - \begin{bmatrix} A \\ B \end{bmatrix} \right\|_F = \left\| \begin{bmatrix} \tilde{C} - C \\ \tilde{S} - S \end{bmatrix} V^H R \right\|_F \leq \epsilon \|R\|_F \quad \blacksquare$$

1998 年, Drmac 提出了计算广义奇异值分解的正切算法 (tangent algorithm) [135]。这种算法分两个阶段进行: 第一阶段将矩阵束 (A, B) 简化为一个矩阵 F ; 第二阶段计算矩阵 F 的奇异值分解。正切算法的理论基础是: 广义奇异值分解在等价变换下是不变的, 即

$$(A, B) \rightarrow (A', B') = (U^T A S, V^T B S) \quad (6.5.18)$$

式中, U, V 是任意的正交矩阵, 且 S 是任意的非奇异矩阵。因此, 根据定义, 两个矩阵束 (A, B) 和 (A', B') 具有相同的广义奇异值分解。

算法 6.5.4 (广义奇异值分解的正切算法) [135]

输入 矩阵 $A = [a_1, a_2, \dots, a_n] \in R^{m \times n}$, $B \in R^{p \times n}$, $m \geq n$, $\text{rank}(B) = n$

步骤 1 计算

$$\begin{aligned} \Delta_A &= \text{diag}(\|a_1\|_2, \|a_2\|_2, \dots, \|a_n\|_2) \\ A_c &= A \Delta_A^{-1}, \quad B_1 = B \Delta_A^{-1} \end{aligned}$$

步骤 2 利用具有列旋转的 Householder QR 分解算法计算

$$\begin{bmatrix} R \\ O \end{bmatrix} = Q^T B_1 \Pi$$

步骤 3 通过求解矩阵方程 $FR = A_c \Pi$, 计算 $F = A_c \Pi R^{-1}$ 。

步骤 4 计算矩阵 F 的奇异值分解

$$\begin{bmatrix} \Sigma \\ O \end{bmatrix} = V^T F U$$

步骤 5 计算矩阵

$$X = \Delta_A^{-1} \Pi R^{-1} U, \quad W = Q \begin{bmatrix} U & O \\ O & I_{p-n} \end{bmatrix}$$

输出 (A, B) 的广义奇异值分解读作

$$\begin{bmatrix} V^T & A \\ W^T & B \end{bmatrix} X = \begin{bmatrix} \Sigma & O \\ I & O \end{bmatrix}$$

6.5.4 二次型不等式约束最小二乘

在最小二乘中,经常需要在 R^n 的一个真子集内极小化 $\|Ax - b\|_2$ 。例如,当我们试图对有噪声的数据拟合一函数时,需要

$$\text{在 } \|Bx\|_2 \leq \alpha \text{ 约束下, 极小化 } \|Ax - b\|_2 \quad (6.5.19)$$

其中, $A \in R^{m \times n}$, $b \in R^m$, $B \in R^{n \times n}$ (非奇异), 且 $\alpha \geq 0$ 。这样一种在二次型不等式约束下的最小二乘问题简称为二次型不等式约束最小二乘 (LSQI) 问题^[184]。

更一般地,有下列问题:

$$\text{在 } \|Bx - d\|_2 \leq \alpha \text{ 约束下, 极小化 } \|Ax - b\|_2 \quad (6.5.20)$$

其中, $A \in R^{m \times n}$ ($m \geq n$), $b \in R^m$, $B \in R^{p \times n}$, $d \in R^p$ 和 $\alpha \geq 0$ 。广义奇异值分解清楚地表明了 LSQI 问题的可解性。若

$$\left. \begin{aligned} U^T A X &= \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n), \quad U^T U = I_m \\ V^T B X &= \text{diag}(\beta_1, \beta_2, \dots, \beta_q), \quad V^T V = I_p, \quad q = \min\{p, n\} \end{aligned} \right\} \quad (6.5.21)$$

是 (A, B) 的广义奇异值分解, 则 LSQI 问题变换成

$$\text{在 } \|D_B y - \tilde{d}\|_2 \leq \alpha \text{ 约束下, 极小化 } \|D_A y - \tilde{b}\|_2$$

其中, $\tilde{b} = U^T b$, $\tilde{d} = V^T d$, $y = X^{-1} x$ 。目标函数的直接形式为

$$\|D_A y - \tilde{b}\|_2^2 = \sum_{i=1}^n (\alpha_i y_i - b_i)^2 + \sum_{i=n+1}^m \tilde{b}_i^2 \quad (6.5.22)$$

约束方程为

$$\|D_B y - \tilde{d}\|_2^2 = \sum_{i=1}^r (\beta_i y_i - \tilde{d}_i)^2 + \sum_{i=r+1}^m \tilde{d}_i^2 \leq \alpha^2 \quad (6.5.23)$$

这就简化了 LSQI 问题的分析。这里, $r = \text{rank}(B)$, 并假定 $\beta_{r+1} = \beta_{r+2} = \dots = \beta_q = 0$ 。

下面分析简化之后的 LSQI 问题式 (6.5.23)。首先, 当且仅当

$$\sum_{i=r+1}^p \tilde{d}_i^2 \leq \alpha^2$$

LSQI 问题有解。如果在上式中取等号, 则式 (6.5.22) 和式 (6.5.23) 一并表明, 由

$$y_i = \begin{cases} \tilde{d}_i/\beta_i, & i = 1, 2, \dots, r \\ \tilde{b}_i/\alpha_i, & i = r+1, r+2, \dots, n; \alpha \neq 0 \\ 0, & i = r+1, r+2, \dots, n; \alpha = 0 \end{cases} \quad (6.5.24)$$

定义的向量是 LSQI 问题的解。另一方面

$$\sum_{i=r+1}^p \tilde{d}_i^2 < \alpha^2 \quad (6.5.25)$$

时, 向量

$$y_i = \begin{cases} \tilde{b}_i/\alpha_i, & \alpha_i \neq 0 \\ \tilde{d}_i/\beta_i, & \alpha_i = 0 \end{cases} \quad i = 1, 2, \dots, n$$

是 $\|D_A y - \tilde{b}\|_2$ 的极小化变量。如果这个向量也是可行的, 则我们就得到了式 (6.5.20) 的解 (然而, 它不一定是最小范数解)。因此, 假定

$$\sum_{i=1, \alpha_i \neq 0} \left(\beta_i \frac{\tilde{b}_i}{\alpha_i} - \tilde{d}_i \right)^2 + \sum_{i=q+1}^p \tilde{d}_i^2 > \alpha^2 \quad (6.5.26)$$

这意味着, LSQI 问题的解发生在可行集的边界。剩下的目标就是

$$\text{在 } \|D_B y - \tilde{d}\|_2 = \alpha \text{ 约束下, 极小化 } \|D_A y - \tilde{b}\|_2$$

可以用 Lagrange 乘数法求解这一问题。为此, 定义

$$Q(\lambda, y) = \|D_A y - \tilde{b}\|_2^2 + \lambda(\|D_B y - \tilde{d}\|_2^2 - \alpha^2)$$

并令 $\partial Q/\partial y_i = 0, i = 1, 2, \dots, n$, 即得线性方程组

$$(D_A^T D_A + \lambda D_B^T D_B) y = D_B^T \tilde{b} + \lambda D_B^T \tilde{d}$$

假定系数矩阵是非奇异的, 则上述方程组的解向量 $y(\lambda)$ 的元素为

$$y_i(\lambda) = \begin{cases} \frac{\alpha_i \tilde{b}_i + \lambda \beta_i \tilde{d}_i}{\alpha_i^2 + \lambda \beta_i^2}, & i = 1, 2, \dots, q \\ \frac{\tilde{b}_i}{\alpha_i}, & i = q+1, q+2, \dots, n \end{cases}$$

式中, Lagrange 乘数 λ 由

$$\phi(\lambda) = \|D_B y(\lambda) - \tilde{d}\|_2^2 = \sum_{i=1}^r \alpha_i \frac{\beta_i \tilde{b}_i - \alpha_i \tilde{d}_i^2}{\alpha_i^2 + \lambda \beta_i^2} + \sum_{i=r+1}^p \tilde{d}_i^2 = \alpha^2$$

的解决定。这类方程称为长期方程 (secular equations)。由式 (6.5.26) 可看出 $\phi(0) > \alpha^2$ 。又由于 $\phi(\lambda)$ 对于 $\lambda > 0$ 是单调递减的, 因此, 式 (6.5.26) 意味着存在唯一的正 λ^* 满足

$\phi(\lambda^*) = \alpha^2$ 。可以利用任一种标准的求根方法求出这个根。最后, 原 LSQI 问题的解由 $\mathbf{x} = \mathbf{X}\mathbf{y}(\lambda^*)$ 给出。

最近几年, 广义奇异值分解在语音信号处理和模式识别中相继被应用。

多麦克风在离散时间 k 采集的含噪声语音信号可以用观测模型

$$\mathbf{y}[k] = \mathbf{x}[k] + \mathbf{v}[k]$$

描述。式中, $\mathbf{x}[k]$ 和 $\mathbf{v}[k]$ 分别为语音信号向量和加性噪声向量。若令 $\mathbf{R}_{yy} = \mathbf{E}\{\mathbf{y}[k]\mathbf{y}^T[k]\}$, $\mathbf{R}_{vv} = \mathbf{E}\{\mathbf{v}[k]\mathbf{v}^T[k]\}$ 分别代表观测数据的自相关矩阵和加性噪声的自相关矩阵, 则可以对它们进行联合对角化, 即

$$\left. \begin{aligned} \mathbf{R}_{yy} &= \mathbf{Q}\text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2)\mathbf{Q}^T \\ \mathbf{R}_{vv} &= \mathbf{Q}\text{diag}(\eta_1^2, \eta_2^2, \dots, \eta_m^2)\mathbf{Q}^T \end{aligned} \right\} \quad (6.5.27)$$

2002 年, Doclo 与 Moonen^[128] 证明了, 为了实现多麦克风语音增强, 使均方误差最小的最优滤波器为

$$\mathbf{W}[k] = \mathbf{R}_{yy}^{-1}[k]\mathbf{R}_{xx}[k] = \mathbf{R}_{yy}^{-1}[k](\mathbf{R}_{yy}[k] - \mathbf{R}_{vv}[k]) \quad (6.5.28)$$

$$= \mathbf{Q}^{-T}\text{diag}\left(1 - \frac{\sigma_1^2}{\eta_1^2}, 1 - \frac{\sigma_2^2}{\eta_2^2}, \dots, 1 - \frac{\sigma_m^2}{\eta_m^2}\right)\mathbf{Q} \quad (6.5.29)$$

构造 $p \times m$ 观测数据矩阵 $\mathbf{Y}[k]$ 和 $q \times m$ 加性噪声矩阵 $\mathbf{V}[k']$ 如下:

$$\mathbf{Y}[k] = \begin{bmatrix} \mathbf{y}^T[k-p+1] \\ \vdots \\ \mathbf{y}^T[k-1] \\ \mathbf{y}^T[k] \end{bmatrix}, \quad \mathbf{V}[k'] = \begin{bmatrix} \mathbf{v}^T[k'-q+1] \\ \vdots \\ \mathbf{v}^T[k'-1] \\ \mathbf{v}^T[k'] \end{bmatrix} \quad (6.5.30)$$

式中, $\mathbf{V}[k']$ 是平时在无语音信号时测量得到的相同环境下的加性噪声数据矩阵。于是, 只要计算矩阵束 $(\mathbf{Y}[k], \mathbf{V}[k'])$ 的广义奇异值分解, 得到 \mathbf{Q} 和广义奇异值 σ_i/η_i , 即可直接获得最优滤波器 $\mathbf{W}^T[k]$ 。理论和仿真结果表明, 这种基于广义奇异值分解的最优滤波器显示了波束形成器的空间指向特性, 有着很好的多麦克风语音增强效果。

在信息恢复系统中, 降维技术对处理大批量数据是至关重要的。为此, 数据的低维表示必须是全部文本数据一个很好的逼近。模式识别通过使类内散布最小、类间散布最大, 对数据进行聚类。然而, 这种识别分析要求类内散布矩阵或类间散布矩阵必须有一个是非奇异的。但是, 文本数据矩阵往往不能满足这一要求。2003 年, Howland 等人^[231] 证明了, 利用广义奇异值分解, 无论文本数据维数多少, 都可以实现聚类; 并且直接使用数据矩阵的广义奇异值分解, 还可避免使用散布矩阵带来的数值稳定性问题。基于广义奇异值分解, 文献^[231] 提出了聚类文本数据的降维方法, 这种方法能够有效保持文本数据的结构。

6.6 约束奇异值分解

前几节分别介绍了奇异值分解 (SVD), 矩阵乘积的奇异值分解 (PSVD) 和矩阵束 (A, B) 的广义奇异值分解 (GSVD), 并在上一节分析了三种奇异值分解之间的关系。

现在考虑一个新的问题: 假定矩阵

$$A_0 = \begin{bmatrix} 0 & 1 \\ a_2 & a_1 \end{bmatrix}$$

是从二阶常微分方程

$$\frac{d^2x}{dt^2} - a_2 \frac{dx}{dt} - a_1 x = f$$

得到的, 则矩阵 A_0 的四个元素中只有 a_1 和 a_2 有误差, 元素“0”和“1”是精确的。于是, 矩阵 $A = A_0 + E$ 的误差矩阵 E 可以有三种情况。

情况 1 仅 a_2 可变

$$E = \begin{bmatrix} 0 & 0 \\ e_{21} & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e_{21} [1 \ 0]$$

情况 2 仅 a_1 可变

$$E = \begin{bmatrix} 0 & 0 \\ 0 & e_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e_{22} [0 \ 1]$$

情况 3 a_1 和 a_2 均可变

$$E = \begin{bmatrix} 0 & 0 \\ e_{21} & e_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} [e_{21} \ e_{22}]$$

观察知, 上述三种形式的任何一种误差矩阵 E 都不可能改变原矩阵 A_0 的秩。然而, 前面几节的三种奇异值分解都不能够得到并解释这一结果。这说明, 需要有另外的奇异值分解方法来分析这类问题。本节介绍解决这类问题的约束奇异值分解方法。约束奇异值分解略写为 RSVD (restricted SVD), 它是由 Zha^[499] 提出, 并由 Moor 与 Golub^[111] 进一步分析的。

6.6.1 约束奇异值

在广义奇异值中, 我们考虑的是矩阵束 (A, B) 的奇异值。正如情况 1 至情况 3 所示, 误差矩阵一般可表示为三个矩阵的乘积, 即 $E = BDC$ 的形式。如果要考虑复矩阵 $A = A_0 + E$ 的秩的确定, 很自然地会使我们联想到应该考虑矩阵三元组 (A, B, C) 的奇异值。令 $A \in C^{m \times n}$, $B \in C^{m \times p}$, $D \in C^{p \times q}$, $C \in C^{q \times n}$, 则有以下定义。

定义 6.6.1 (约束奇异值)^[499] 矩阵三元组 (A, B, C) 的约束奇异值定义为

$$\sigma_k(A, B, C) = \min_{D \in C^{p \times q}} \{\|D\|_F | \text{rank}(A + BDC) \leq k-1\}, \quad k = 1, 2, \dots, n \quad (6.6.1)$$

关于上述定义, 应注意以下几点。

注释1 如果对于某个 $k(1 \leq k \leq n)$, 不存在满足 $\text{rank}(A + BDC) \leq k-1$ 的矩阵 D , 则约束奇异值 $\sigma_k(A, B, C)$ 定义为无穷大。

注释2 如果 $m < n$ 和 $m < k \leq n$, 则 $\sigma_k(A, B, C) = 0$ 。

注释3 从定义容易看出, 约束奇异值按非减次序排列, 即

$$\sigma_k(A, B, C) \geq \sigma_{k+1}(A, B, C), \quad k = 1, 2, \dots, n-1$$

作为进一步的说明, 仍然以情况 1~ 情况 3 为例, 它们的约束奇异值分别如下 [499]。

对于情况 1

$$A = \begin{bmatrix} 0 & 1 \\ a_2 & a_1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [1, 0]$$

约束奇异值为

$$\sigma_1(A, B, C) = \infty, \quad \sigma_2(A, B, C) = |a_2|$$

对于情况 2

$$A = \begin{bmatrix} 0 & 1 \\ a_2 & a_1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [0, 1]$$

约束奇异值为

$$\sigma_1(A, B, C) = \infty, \quad \sigma_2(A, B, C) = \infty, \quad \text{若 } a_2 \neq 0$$

$$\sigma_1(A, B, C) = \infty, \quad \sigma_2(A, B, C) = 0, \quad \text{若 } a_2 = 0$$

对于情况 3

$$A = \begin{bmatrix} 0 & 1 \\ a_2 & a_1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

约束奇异值为

$$\sigma_1(A, B, C) = \infty, \quad \sigma_2(A, B, C) = |a_2|$$

以上是低阶矩阵三元组的情况, 它们的约束奇异值容易直接观察得出。对于高阶的矩阵三元组, 约束奇异值需要有约束奇异值分解方法才能求出, 详见下一小节讨论。

有趣的是, 前面几节介绍的奇异值和广义奇异值都可以看作是约束奇异值的特例。

(1) 矩阵 A 的奇异值可以通过令 $B = I_m$ 和 $C = I_n$ 从矩阵三元组 (A, B, C) 的约束奇异值得到。

(2) 矩阵束 (A, C) 的广义奇异值就是矩阵三元组 (A, I_m, C) 的约束奇异值。

更严格地, 我们可以用下面的两个命题来叙述以上结论。

命题 6.6.1 [499] 假定矩阵 $B \neq O$ 和 $C \neq O$, 则

$$\sigma_k(A, I_m, I_n) = \sigma_k, \quad k = 1, 2, \dots, n \quad (6.6:2)$$

式中, σ_k 是矩阵 A 的奇异值。

命题 6.6.2^[499] 约束奇异值和广义奇异值之间存在下列关系:

$$\sigma_i(A, I_m, C) = \begin{cases} \frac{\alpha_i}{\beta_i}, & i = 1, 2, \dots, k \\ 0, & i = k+1, k+2, \dots, n \end{cases} \quad (6.6.3)$$

6.6.2 约束奇异值分解

上一小节定义了约束奇异值,并分析了约束奇异值与奇异值、广义奇异值之间的关系。现在,我们来讨论如何实现约束奇异值分解,以得到约束奇异值。我们假定矩阵 B 和 C 为一般的矩阵。实现约束奇异值分解的关键是下面的重要观察。

引理 6.6.1 令 $P \in C^{m \times m}$ 和 $Q \in C^{n \times n}$ 是非奇异的矩阵,并令 $U \in C^{p \times p}$ 和 $V \in C^{q \times q}$ 为酉矩阵,则

$$\sigma_k(PAQ, PBU, VCQ) = \sigma_k(A, B, C), \quad k = 1, 2, \dots, n \quad (6.6.4)$$

上述引理规定了对矩阵 A, B, C 进行的线性变换的类型,它们必须能够保持矩阵三元组的约束奇异值不改变。因此,约束奇异值分解的关键就是寻找这样的变换。

定理 6.6.1 (约束奇异值分解)^[499] 令 $A \in C^{m \times n}$, $B \in C^{m \times p}$, $C \in C^{q \times n}$, 则存在非奇异矩阵 $P \in C^{m \times m}$ 和 $Q \in C^{n \times n}$ 以及酉矩阵 $U \in C^{p \times p}$ 和 $V \in C^{q \times q}$ 使得

$$PAQ = \begin{matrix} n-t_1 & t_1 \\ \left[\begin{array}{cc} \Sigma_A & \\ & O_A^{(2)} \end{array} \right] & \begin{matrix} m-t_1 \\ t_1 \end{matrix} \end{matrix}, \quad PBU = \begin{matrix} n-t_1 & t_1 \\ \left[\begin{array}{cc} \Sigma_B & \\ & O_B^{(2)} \end{array} \right] & \begin{matrix} m \\ m-t_2 \end{matrix} \end{matrix}, \quad VCQ = \begin{matrix} n-t_1 & t_1 \\ \left[\begin{array}{cc} \Sigma_C & \\ & O_C^{(2)} \end{array} \right] & \end{matrix}$$

成立,其中

$$\Sigma_A = \begin{matrix} & j & k & l & r & s_1 \\ \left[\begin{array}{cccccc} I_j & & & & & \\ & I_k & & & & \\ & & I_l & & & \\ & & & S_A & & \\ & & & & O_A^{(1)} & \end{array} \right] & \begin{matrix} j \\ k \\ l \\ r \\ s_2 \end{matrix} \end{matrix} \quad (6.6.5)$$

$$\Sigma_B = \begin{matrix} & j & p-j-r-s_2 & r & s_2 \\ \left[\begin{array}{cccc} I_j & & & \\ & O_B^{(1)} & & \\ & & S_B & \\ & & & I_{s_2} \end{array} \right] & \begin{matrix} j \\ k+l \\ r \\ s_2 \end{matrix} \end{matrix} \quad (6.6.6)$$

$$\Sigma_C = \begin{matrix} & j+k & l & r & s_1 \\ \left[\begin{array}{cccc} O_C^{(1)} & & & \\ & I_l & & \\ & & S_C & \\ & & & I_{s_1} \end{array} \right] & \begin{matrix} q-l-r-s_1 \\ l \\ r \\ s_1 \end{matrix} \end{matrix} \quad (6.6.7)$$

式中, $S_A = \text{diag}(\alpha_i)$, $S_B = \text{diag}(\beta_i)$, $S_C = \text{diag}(\gamma_i)$, 并且

$$\alpha_i^2 + \beta_i^2 + \gamma_i^2 = 1, \quad i = s+1, s+2, \dots, s+r \quad (6.6.8)$$

其中

$$1 > \alpha_i \geq \alpha_{i+1} > 0, \quad 0 < \beta_i \leq \beta_{i+1} < 1, \quad 1 > \gamma_i \geq \gamma_{i+1} > 0 \quad (6.6.9)$$

和

$$\frac{\alpha_i}{\beta_i \gamma_i} \geq \frac{\alpha_{i+1}}{\beta_{i+1} \gamma_{i+1}}, \quad i = s+1, s+2, \dots, s+r-1 \quad (6.6.10)$$

其中, $s = j + k + l$ 。

证明 参见文献 [499]。

类似于广义奇异值, 定义

$$\alpha_i = 1, \quad \beta_i = 1, \quad \gamma_i = 0, \quad i = 1, 2, \dots, j$$

$$\alpha_i = 1, \quad \beta_i = 0, \quad \gamma_i = 0, \quad i = j+1, j+2, \dots, j+k$$

$$\alpha_i = 1, \quad \beta_i = 0, \quad \gamma_i = 1, \quad i = j+k+1, j+k+2, \dots, s$$

$$\alpha_i, \beta_i, \gamma_i \text{ 取 } S_A, S_B, S_C \text{ 中的值,} \quad i = s+1, s+2, \dots, s+r$$

$$\alpha_i = 0, \quad \beta_i = 1, \quad \gamma_i = 1, \quad i = s+r+1, s+r+2, \dots, s+r+\min\{s_1, s_2\}$$

是矩阵三元组 (A, B, C) 的非平凡约束奇异值。

约束奇异值分解使得可以对矩阵 A 计算一个低秩近似矩阵, 方法是从矩阵 A 减去一个矩阵, 被减去的矩阵必须服从约束条件: 它的列位于 B 的列空间, 而行则位于 C 的行空间。

有必要指出, 对于矩阵 A, B 和 C 的某些特殊选择, 约束奇异值分解 (RSVD) 将简化为普通奇异值分解 (OSVD)、乘积奇异值分解 (PSVD) 和广义奇异值分解 (GSVD)。下面的定理严格地叙述了这些结果。

定理 6.6.2 约束奇异值分解的特殊情况包括^[111]

- (1) 矩阵三元组 (A, I_m, I_n) 的 RSVD 就是矩阵 A 的 OSVD;
- (2) 矩阵三元组 (I_m, B, C) 的 RSVD 就是 (B^H, C) 的 PSVD;
- (3) 矩阵三元组 (A, B, I_n) 的 RSVD 就是矩阵束 (A, B) 的 GSVD;
- (4) 矩阵三元组 (A, I_m, C) 的 RSVD 就是矩阵束 (A, C) 的 GSVD。

证明^[111] 下面分别证明之。

(1) 令 $B = I_m, C = I_n$, 并考虑矩阵三元组 (A, B, C) 的 RSVD。根据定理 6.6.1 有 $I = P^{-1}S_B U^H$ 和 $I_n = V S_C Q^{-1}$ 。这意味着, $P = S_B U^H$ 和 $Q = V S_C$ 。因此, 可以求得

$$A = V_B (S_B^{-1} S_A S_C^{-1}) U^H$$

这恰好就是矩阵 A 的 OSVD。

(2) 令 $A = I_m$, 并考虑 (I_m, B, C) 的 RSVD。于是, $I_m = P^{-1}S_A Q^{-1}$, 这意味着 $Q^{-1} = S_A^{-1}P$ 。因此, $B^H = V_B S_B^T P^{-1}, C = U_C (S_C S_1^{-1}) P^H$, 这正是 (B^H, C) 的 PSVD。

(3) 令 $C = I_n$, 并考虑 (A, B, I_n) 对称矩阵 RSVD。于是, $I_n = U_C S_C Q^{-1}$, 从而 $Q^{-1} = S_C^{-1} U_C^H$ 。因此, $A = P^{-1}(S_A S_C^{-1}) U_C^H$ 和 $B = P^{-1} S_B V_B^H$, 这就是矩阵束 (A, B) 的 GSVD(相差一个对角线比例因子)。

(4) 令 $B = I_m$ 后, 证明与 (3) 类似。 ■

6.7 结构奇异值

作为具有结构不确定性的反馈系统的分析与综合的一种工具, Doyle 于 1982 年提出了结构奇异值 (structured singular value) 的概念^[132]。结构奇异值在 H^∞ 控制和系统特征的描述中有着重要的应用^{[133], [150]}。本节介绍结构奇异值的几种定义以及如何将结构奇异值的计算转换成几个光滑最优化问题的求解。

6.7.1 结构奇异值的定义与性质

对于一个正方的复矩阵 M , 用 $\rho(M)$ 表示其谱半径, 用 $\bar{\sigma}(M)$ 表示其最大的奇异值。

任何一个由正整数组成的 m 元组 $\mathcal{K} = (k_1, k_2, \dots, k_m)$ 称作尺度为 m 的分块结构。给定一个尺度为 m 的分块结构 \mathcal{K} , 定义投影矩阵为

$$P_i = \text{diag}(O_{k_1}, \dots, O_{k_{i-1}}, I_{k_i}, O_{k_{i+1}}, \dots, O_{k_m})$$

块对角矩阵族为

$$\mathcal{D} = \{\text{diag}(d_1 I_{k_1}, \dots, d_m I_{k_m}) \mid d_i \in (0, \infty)\}$$

和块酉矩阵族为

$$\mathcal{U} = \{\text{diag}(U_1, U_2, \dots, U_m) \mid U_i \text{ 是 } k_i \times k_i \text{ 酉矩阵}\}$$

对于任意正标量 δ (有可能是 ∞), 我们记块对角矩阵族

$$\mathcal{X}_\delta = \{\text{diag}(\Delta_1, \Delta_2, \dots, \Delta_m) \mid \Delta_i \text{ 是满足条件 } \bar{\sigma}(\Delta_i) \leq \delta \text{ 的 } k_i \times k_i \text{ 复矩阵}\}$$

所有以上矩阵都具有维数 $n \times n$, 其中

$$n = \sum_{j=1}^m k_j$$

定义 6.7.1 (结构奇异值)^[132] 一个复值的 $n \times n$ 矩阵 M 相对于分块结构 \mathcal{K} 的结构奇异值是具有以下性质的正数 μ :

$$\det(I + M\Delta) \neq 0, \quad \forall \Delta \in \mathcal{X}_\delta$$

当且仅当

$$\mu\delta < 1$$

换言之, 如果在 X_∞ 内没有满足 $\det(I + M\Delta) = 0$ 的 Δ 存在, 则 $\mu(M) = 0$; 否则, 有 $\mu(M) = (\min_{\Delta \in X_\infty} \{\bar{\sigma}(\Delta) \mid \det(I + M\Delta) = 0\})^{-1}$.

应当注意, $P_i, \mathcal{D}, \mathcal{U}, X_\delta$ 和 $\mu(M)$ 均与分块结构有关。不过, 在大多数情况下, 并不明显规定这种分块结构。

关于结构奇异值, 容易验证以下事实^[132]。

事实 1 对于所有 $U \in \mathcal{U}$, 有

$$\mu(M) = \mu(MU) = \mu(UM)$$

并且对于所有 $D \in \mathcal{D}$, 有

$$\mu(M) = \mu(DMD^{-1})$$

事实 2

$$\mu(M) = \max_{U \in \mathcal{U}} \rho(MU) = \max_{U \in \mathcal{U}} \rho(UM) \quad (6.7.1)$$

事实 3

$$\mu(M) \leq \inf_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}) \quad (6.7.2)$$

并且对于尺度不超过 3 的分块结构, 有

$$\mu(M) = \inf_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}) \quad (6.7.3)$$

其中, \inf 代表下确界。

事实 2 和 3 提供了比定义 6.7.1 更适合于计算的结构奇异值定义。然而, 下面的定义将结构奇异值同光滑 (smooth) 最优化问题建立了联系, 是一种更容易操作的定义。

定义 6.7.2 一个复值的 $n \times n$ 矩阵 M 相对于分块结构 \mathcal{K} 的结构奇异值定义为

$$\mu(M) = \max_{\mathbf{x} \in C^n} \{\|M\mathbf{x}\| : \|P_i\mathbf{x}\|\|M\mathbf{x}\| = \|P_iM\mathbf{x}\|, i = 1, 2, \dots, m\} \quad (6.7.4)$$

或

$$\mu(M) = \max_{\mathbf{x} \in C^n} \{\|M\mathbf{x}\| : \|P_i\mathbf{x}\|\|M\mathbf{x}\| \leq \|P_iM\mathbf{x}\|, i = 1, 2, \dots, m\} \quad (6.7.5)$$

上述定义是由 Fan 与 Tits^[150] 以定理形式证明的结果。他们还证明了, 如果极大化被约束在单位球 $\{\mathbf{x} \in C^n \mid \|\mathbf{x}\| \leq 1\}$ 内或单位球面 $\{\mathbf{x} \in C^n \mid \|\mathbf{x}\| = 1\}$ 上, 定义中的两个等式仍然成立。

式 (6.7.4) 和式 (6.7.5) 的极大化结果显然不是唯一的, 因为如果 \mathbf{x}^* 是一个极大化结果的话, 那么对于任何幅值为 1 的复数 θ 而言, 则 $\theta\mathbf{x}^*$ 也是一个极大化结果。

与结构奇异值定义式 (6.7.1) 和式 (6.7.3) 相比, 式 (6.7.4) 和式 (6.7.5) 中的目标函数和约束条件更方便计算; 在对所有的范数取平方后, 式 (6.7.4) 和式 (6.7.5) 中的目标函数和约束条件是光滑的。

6.7.2 结构奇异值的计算

现在考虑结构奇异值的计算方法。给出任一矩阵 M 和分块结构 K , 下面的算法可以构造一个简化矩阵 \bar{M} , 它具有与 M 相同的秩, 且相应的简化分块结构 \bar{K} 的任一尺度都不会大于 M 的秩。

算法 6.7.1 (简化矩阵的构造算法)^[150]

输入

$$M \in C^{n \times n}, \quad K = (k_1, k_2, \dots, k_m)$$

步骤 1 实现矩阵 M 的奇异值分解。令 $r = \text{rank}(M)$, 并令 U_i 和 V_i 均是 $k_i \times r$ 矩阵, 其中, $i = 1, 2, \dots, m$ 。又令

$$\begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_m \end{bmatrix} \quad \text{和} \quad \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix}$$

分别由 M 的前 r 个左奇异向量和右奇异向量组成。于是, 对于某个正定的对角矩阵 $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, 有

$$M = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_m \end{bmatrix} \Sigma \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix}^H$$

步骤 2 计算 $W^U = \text{diag}(W_1^U, W_2^U, \dots, W_m^U)$ 和 $W^V = \text{diag}(W_1^V, W_2^V, \dots, W_m^V)$, 这两个矩阵都在块酉矩阵族 \mathcal{U} 内, 且对于 $i = 1, 2, \dots, m$, 矩阵 $W_i^U U_i$ 的最后 $(k_i - r_i^U)$ 行和 $W_i^V V_i$ 的最后 $(k_i - r_i^V)$ 行均为零, 其中, r_i^U 和 r_i^V 分别是 U_i 和 V_i 的秩。

步骤 3 对于 $i = 1, 2, \dots, m$, 令 $\bar{k}_i = \max(r_i^U, r_i^V)$, 并令 \bar{U}_i 和 \bar{V}_i 分别由 $W_i^U U_i$ 和 $W_i^V V_i$ 的前 \bar{k}_i 行组成, 则

$$\bar{M} = \begin{bmatrix} \bar{U}_1 \\ \bar{U}_2 \\ \vdots \\ \bar{U}_m \end{bmatrix} \Sigma \begin{bmatrix} \bar{V}_1 \\ \bar{V}_2 \\ \vdots \\ \bar{V}_m \end{bmatrix}^H$$

和

$$\bar{K} = (\bar{k}_1, \bar{k}_2, \dots, \bar{k}_m)$$

应当注意的是, 矩阵 W^U 和 W^V 都不是唯一的。因此, 只能够把 \bar{M} 视为对应于 M 的一个简化矩阵。将矩阵 M 简化为 \bar{M} 的运算可以看做是对 M 进行一系列矩阵变换。如下面的命题所述, 结构奇异值在这些变换下可保持不变。

命题 6.7.1^[150] 由算法 6.7.1 构造的简化矩阵 \bar{M} 和分块结构 \bar{K} 满足关系

$$\mu(M) = \bar{\mu}(\bar{M})$$

式中, $\bar{\mu}$ 表示相对于分块结构 \bar{K} 的结构奇异值。

作为算法 6.7.1 的一个特例, 下面的定理给出了秩等于 1 的矩阵 M 的结构奇异值的计算方法。

定理 6.7.1 (秩 1 矩阵的结构奇异值)^[150] 假定矩阵 M 具有秩 1, 即 $M = uv^H$ 对 $u, v \in C^n$ 。令 $\bar{u}, \bar{v} \in R^m$ 定义为

$$\begin{aligned}\bar{u} &= [\bar{u}_1, \bar{u}_2, \dots, \bar{u}_m]^T \\ \bar{v} &= [\bar{v}_1, \bar{v}_2, \dots, \bar{v}_m]^T\end{aligned}$$

其中, $\bar{u}_i = \|P_i u\|, \bar{v}_i = \|P_i v\|, i = 1, 2, \dots, m$, 则 $\mu(M) = \bar{u}^T \bar{v}$ 。

将式 (6.7.5) 视为一个在实向量空间 R^{2n} 的最优化问题。令 $z = \begin{bmatrix} \text{Re } x \\ \text{Im } x \end{bmatrix} \in R^{2n}$, 则式 (6.7.5) 中关于 z 的约束条件可以写作

$$g_i(z) = \|P_i x\|^2 \|Mx\|^2 - \|P_i Mx\|^2 \leq 0, \quad i = 1, 2, \dots, m$$

定义 6.7.3 (正则点)^[150] 对于任意 $x \in C^n$, 若与之对应的 $z \in R^{2n}$ 使得 $\{\nabla g_i(z), i = 1, 2, \dots, m\}$ 组成一个在整个实数域都线性无关的向量集合, 则称该 x 是式 (6.7.4) 的一个正则点。

对于一般的矩阵, 其结构奇异值与最优化问题之间存在下列关系。

定理 6.7.2^[150] 考虑下面三个最优化问题 (定义 6.7.2, 事实 3, 事实 2):

- (P1) $\mu(M) = \max_{x \in C^n} \{\|Mx\| : \|P_i x\| \|Mx\| = \|P_i Mx\|, i = 1, 2, \dots, m\}$
 (P2) $\inf_{D \in \mathcal{D}} \delta(DMD^{-1})$
 (P3) $\max_{U \in \mathcal{U}} \rho(MU)$

并假定: ① x^* 是形如 θx^* 的优化问题 (P1) 的唯一全局解, 其中, θ 是一个单位幅值的复数, 而 x^* 是 (P1) 的正则点; ② 式 (6.7.3) 成立; ③ (P2) 中的下确界 “inf” 能够达到。在这些条件下, 以下结果成立:

- (1) 假定 $D^* \in \mathcal{D}$ 是问题 (P2) 的解, 则存在一个与 $D^* M D^{*-1}$ 的最大奇异值相对应的右奇异向量 (例如 y) 使得 $x = (D^{*-1} y^*) / \|D^{*-1} y^*\|$ 是问题 (P1) 的解。
- (2) 令 $\lambda^* = [\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*]^T$ 是与 x^* 相关联的 (唯一) 乘数向量, 则块对角矩阵 $D = \text{diag}(\lambda_i^{*1/2} I_{k_i})$ 属于 \mathcal{D} , 并且 D 是问题 (P2) 的唯一解 (除了相差一个标量乘数外)。
- (3) 对于 $i = 1, 2, \dots, m$, 令 Q_i 为 $k_i \times n$ 矩阵, 它由 P_i 从 $\sum_{j=1}^{i-1} k_j + 1$ 到 $\sum_{j=1}^i k_j$ 的所有行组成。也就是说, Q_i 是 P_i 的第 i 个行分块。令 $U = \text{diag}(U_1, U_2, \dots, U_m)$, 式中, $U_i (i = 1, 2, \dots, m)$ 是 $k_i \times k_i$ 矩阵, 并满足 $U_i = V_i W_i^H$, 其中, V_i 和 W_i 是两个酉矩阵, 它们的第 1 列分别为 $Q_i x^* / \|Q_i x^*\|$ 和 $Q_i M x^* / \|Q_i M x^*\|$, 若 $Q_i x^* \neq 0$ 。此时, U 是问题 (P3) 的解。
- (4) 假定 $U^* \in \mathcal{U}$ 是问题 (P3) 的解, 则对于矩阵 MU^* 与幅值等于谱半径的特征值相对应的任一单位范数的特征向量 z^* , 向量 $x = U^* z^*$ 是问题 (P1) 的解。

现在假定定理 6.7.2 中的假设条件都满足, 并假定已得到问题 (P1) 的一个候补解 \mathbf{x}^* 或问题 (P3) 的一个候补解 \mathbf{U}^* , 但是不能够肯定 $\mu(M)$ 是否已获得, 也就是说, 不能够肯定 \mathbf{x}^* 或 \mathbf{U}^* 是否就是全局极大化结果。下面的算法提供了一种计算结构奇异值 $\mu(M)$ 的合理方法。

算法 6.7.2 (结构奇异值计算算法) ^[150]

步骤 1 如果 M 具有秩 1, 则按照定理 6.7.1 求 $\mu(M)$, 算法即告完成。若对某个 i , M 的秩小于 k_i , 就利用算法 6.7.1 计算 M 的简化矩阵 \bar{M} , 并再次用 M 和 K 表记简化后的矩阵和分块结构。

步骤 2 求解问题 (P1), 得到 (可能是局部的) 极大化变量 $\hat{\mathbf{x}}$ 。

步骤 3 按照定理 6.7.2 中的 (2), 由 $\hat{\mathbf{x}}$ 计算 \hat{D} 。如果 $\|M\hat{\mathbf{x}}\| = \bar{\sigma}(\hat{D}M\hat{D}^{-1})$, 则取 $\mu(M) = \|M\hat{\mathbf{x}}\|$, 算法属成功完成¹。

步骤 4 求解 (P2) 问题, 得到 D^* , 它是全局极小化变量²。

步骤 5 如果 $\max(\rho(M), \|M\hat{\mathbf{x}}\|) = \bar{\sigma}(D^*MD^{*-1})$, 则取 $\mu(M) = \bar{\sigma}(D^*MD^{*-1})$, 算法成功完成³。否则, $\mu(M)$ 的搜寻失败。

显然, 如果定理 6.7.2 的条件均满足, 而且算法 6.7.2 成功完成时, 算法将给出正确的结构奇异值。

前面几节介绍了奇异值分解及其各种推广。最后再叙述一种奇异值分解。它是 Takagi 早在 1925 年提出的对称复矩阵的奇异值分解^[439], 现在习惯上称之为 Takagi 奇异值分解。

定理 6.7.3 (Takagi 奇异值分解) 若 A 是一个对称复矩阵, 则存在一个酉矩阵 U 和一个非负对角矩阵 $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ 使得

$$A = U\Sigma U^H$$

其中, U 的列是矩阵 AA^H 的特征向量的正交集, 并且与之对应的 Σ 的对角线元素是 AA^H 的对应特征值的非负平方根。

6.8 奇异值分解的应用

前面几节定义了奇异值分解以及各种推广的奇异值分解, 并分析了它们的性质。从这一节开始, 我们介绍它们的一些重要应用。

奇异值分解已广泛应用于许多工程问题的解决中。例如, 仅奇异值分解与信号处理的国际学术专题讨论会的论文集就有多部 (例如文献 [319], 文献 [183] 等)。本节选择系统辨识和信号处理中的几个典型例子介绍奇异值分解的应用。

1 如果 \hat{D} 内的任一对角元素为 0, 则用 ϵ 代替, 并使用 $\hat{D}(\epsilon)M\hat{D}(\epsilon)^{-1}$ 代替 $\hat{D}M\hat{D}^{-1}$ 。

2 这里, 0 和 ∞ 都被看做是 D^* 内元素的可接受的值。

3 如果 D^* 内的任何元素为无穷大, 则用 $1/\epsilon$ 代替。还需要注意, 若式 (6.7.3) 成立, 但定理 6.7.2 中的其他条件之一不满足时, 有 $\|M\hat{\mathbf{x}}\| = \bar{\sigma}(D^*MD^{*-1})$ 。

6.8.1 静态系统的奇异值分解

以电子器件为例,我们来考虑静态系统的奇异值分解。假定某电子器件的电压 v 和电流 i 之间存在下列关系 (即静态系统模型为):

$$\underbrace{\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}}_{\mathbf{F}} \begin{bmatrix} v_1 \\ v_2 \\ i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (6.8.1)$$

矩阵 \mathbf{F} 的元素限定取 v_1, v_2, i_1, i_2 的允许值。

如果所用的电压和电流测量装置具有相同的精度 (比如 1%), 那么我们就可以很容易检测任何一组测量值是或不是式 (6.8.1) 在期望的精度范围内的解。假定用各种方法得到另外一个矩阵表达式:

$$\begin{bmatrix} 1 & -1 & 10^6 & 10^6 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (6.8.2)$$

显然, 只有当电流非常精确测量时, 一组 v_1, v_2, i_1, i_2 测量值才会以合适的精度满足式 (6.8.2); 而对于电流测量有 1% 的测量误差的一般情况, 式 (6.8.2) 与静态系统模型 (6.8.1) 是大相径庭的: 式 (6.8.1) 给出的电压关系为 $v_1 - v_2 = 0$, 而由于 $i_1 + i_2 = 0.01$ 的测量误差, 式 (6.8.2) 给出的电压关系则是 $v_1 - v_2 + 10^4 = 0$ 。然而, 从代数的角度看, 式 (6.8.1) 和式 (6.8.2) 是完全等价的。因此, 我们希望能够有某些手段来比较几种代数等价的模型表示, 以确定哪一个是我们所希望的、适用一般而不是特殊情况的通用静态系统模型。解决这个问题的基本数学工具就是奇异值分解。

更一般地, 我们考虑 n 个电阻的静态系统方程^[88]:

$$\mathbf{F} \begin{bmatrix} v \\ i \end{bmatrix} = 0 \quad (6.8.3)$$

式中, \mathbf{F} 是一个 $m \times n$ 矩阵。为了简化表示, 我们将一些不变的补偿项撤去了。这样一种表达式是非常通用的, 它可以来自某些物理装置 (例如线性化的物理方程) 和网络方程。矩阵 \mathbf{F} 对数据的精确部分和非精确部分的作用可以利用奇异值分解来进行分析。令 \mathbf{F} 的奇异值分解为

$$\mathbf{F} = \mathbf{U}^T \mathbf{\Sigma} \mathbf{V} \quad (6.8.4)$$

于是, 精确部分和非精确部分的各个分量被矩阵 \mathbf{F} 的奇异值 $\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0$ 作不同的大小改变。如果式 (6.8.3) 是物理装置设计的准确规格, 那么矩阵 \mathbf{F} 的奇异值分解将提供一个代数等价, 但在数值上却最可靠的设计方程。注意到 \mathbf{U} 是一正交矩阵, 所以由式 (6.8.3) 和式 (6.8.4) 有

$$\mathbf{\Sigma} \mathbf{V} \begin{bmatrix} v \\ i \end{bmatrix} = 0 \quad (6.8.5)$$

若将对角矩阵 Σ 分块为

$$\Sigma = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix}$$

并将正交矩阵 V 作相应的分块, 即

$$V = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

其中, $[A, B]$ 是 V 最上面的 r 行, 则式 (6.8.5) 可以写作

$$\begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} v \\ i \end{bmatrix} = 0$$

从而, 我们可以得到与式 (6.8.3) 在代数上等价, 但在数值上却最可靠的表达式:

$$[A, B] \begin{bmatrix} v \\ i \end{bmatrix} = 0 \quad (6.8.6)$$

如果式 (6.8.3) 是物理装置的不精确模型, 则对角矩阵的对角线上就不会出现零奇异值。这时, 我们就不能够直接使用式 (6.8.6)。在这种情况下, 我们需要对模型进行修正, 方法是令所有奇异值 $\sigma_s, \sigma_{s+1}, \dots$ 等于零, 其中, s 是满足 σ_s/σ_1 小于矩阵 F 的元素所允许的精确度 (即物理装置的测量精确度) 的最小整数。于是, 式 (6.8.6) 中的 $[A, B]$ 修正为 V 的最上面 $s-1$ 行。有关结果表明, 这样一种修正可以使参数的变化限制在预先设定的误差范围内^[88]。

现在考虑一个电阻性的多端对 (电阻、电导、混合参数、传导和散射等) 的不同表达式, 目的是寻找一个尽可能最优的表达式。例如, 使用端对坐标 x 和 y 时, 电阻性多端对的显式表示则为^[88]

$$y = Ax, \quad \begin{bmatrix} y \\ x \end{bmatrix} = \Omega \begin{bmatrix} v \\ i \end{bmatrix} \quad (6.8.7)$$

通过选择合适的坐标变换 Ω , 就可以得到电阻、电导、任意混合参数或传导的表达式。于是, 矩阵 A 的条件数就代表从 x 到 y 的信噪比放大倍数的上限。如果 A 可逆, 则该条件数也是从 y 到 x 的信噪比放大倍数的上限。因此, 不同的表达式就可以根据它们的条件数进行排队。这就使得所有参数化表达式一目了然。显然, 最优的情况是条件数 $\text{cond}(A) = 1$ 或 A 是一正交矩阵 (包含一比例因子)。一个自然会问的问题是, 任何一个多端对的电阻器是否有一个最优的表达式? 也就是说, 是否存在使得 $\text{cond}(A) = 1$ 的正交矩阵 A ? 为此, 让我们来看一个 n 维 n 端对的电阻器的隐含表达式:

$$F \begin{bmatrix} v \\ i \end{bmatrix} = 0, \quad \text{rank}(F) = n \quad (6.8.8)$$

应用 F 的奇异值分解式 (6.8.4), 即可得到式 (6.8.6), 其中, $r = n$ 。选择正交坐标变换

$$\begin{bmatrix} y \\ x \end{bmatrix} = \underbrace{\begin{bmatrix} I/\sqrt{2} & I/\sqrt{2} \\ -I/\sqrt{2} & I/\sqrt{2} \end{bmatrix}}_{\Omega} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} v \\ i \end{bmatrix} \quad (6.8.9)$$

这样一来, 就可以利用 Ω 的正交性 $\Omega^{-1} = Q^T$, 将隐含表达式 (6.8.6) 表示成

$$\begin{aligned} [A, B] \begin{bmatrix} v \\ i \end{bmatrix} &= [A, B] \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} \begin{bmatrix} I/\sqrt{2} & -I/\sqrt{2} \\ I/\sqrt{2} & I/\sqrt{2} \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} \\ &= [I, O] \begin{bmatrix} I/\sqrt{2} & -I/\sqrt{2} \\ I/\sqrt{2} & I/\sqrt{2} \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = 0 \end{aligned}$$

即有

$$\begin{bmatrix} I/\sqrt{2} & -I/\sqrt{2} \\ I/\sqrt{2} & I/\sqrt{2} \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = 0 \Rightarrow y = x \quad (6.8.10)$$

于是, 可以得出结论: 利用式 (6.8.4) 的奇异值分解可以得到式 (6.8.9) 的正交变换, 而通过此正交变换, 即可得到一个在数值上最优的显式关系 $y = x$ 。

6.8.2 系统辨识

考虑时不变、线性离散时间的多变量系统, 它具有状态空间表示:

$$\left. \begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \right\} \quad (6.8.11)$$

式中, $u(k)$, $y(k)$ 和 $x(k)$ 分别表示在时间 k 的输入向量, 输出向量和状态向量, 且向量 $x(k)$ 的维数是最小的系统阶数 n ; A, B, C 和 D 是待辨识的未知系统矩阵。系统辨识问题的提法是, 只利用观测到的输入/输出数据 u_k, u_{k+1}, \dots 和 y_k, y_{k+1}, \dots 辨识系统矩阵 A, B, C 和 D 。

Moonen 等人业已证明^[319], 状态向量序列可以仅根据输入/输出观测值计算如下: 令 H_1 和 H_2 定义为

$$H_1 = \begin{bmatrix} u_k & u_{k+1} & \cdots & u_{k+j-1} \\ y_k & y_{k+1} & \cdots & y_{k+j-1} \\ u_{k+1} & u_{k+2} & \cdots & u_{k+j} \\ y_{k+1} & y_{k+2} & \cdots & y_{k+j} \\ \vdots & \vdots & & \vdots \\ u_{k+i-1} & u_{k+i} & \cdots & u_{k+j+i-2} \\ y_{k+i-1} & y_{k+i} & \cdots & y_{k+j+i-2} \end{bmatrix} \quad (6.8.12)$$

$$H_2 = \begin{bmatrix} u_{k+i} & u_{k+i+1} & \cdots & u_{k+i+j-1} \\ y_{k+i} & y_{k+i+1} & \cdots & y_{k+i+j-1} \\ u_{k+i+1} & u_{k+i+2} & \cdots & u_{k+i+j} \\ y_{k+i+1} & y_{k+i+2} & \cdots & y_{k+i+j} \\ \vdots & \vdots & & \vdots \\ u_{k+2i-1} & u_{k+2i} & \cdots & u_{k+2i+j-2} \\ y_{k+2i-1} & y_{k+2i} & \cdots & y_{k+2i+j-2} \end{bmatrix} \quad (6.8.13)$$

其中, $j \geq 2(m+l)i$, 并定义状态向量序列

$$x = [x_{k+i}, x_{k+i+1}, \dots, x_{k+i+j-1}]$$

则在一定条件下, 有

$$\text{Span}_{\text{row}}(\mathbf{x}) = \text{Span}_{\text{row}}(\mathbf{H}_1) \cap \text{Span}_{\text{row}}(\mathbf{H}_2)$$

因此, 此交集的任何基都组成一个合适的状态向量序列 \mathbf{x} , 它具有与相邻行向量相同的基向量。

一旦 $\mathbf{x} = [x_{k+i}, x_{k+i+1}, \dots, x_{k+i+j-1}]$ 已知, 则系统矩阵可以通过求解 (超定的) 线性方程组

$$\begin{bmatrix} x_{k+i+1} & \cdots & x_{k+i+j-1} \\ y_{k+i} & \cdots & y_{k+i+j-2} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} x_{k+i} & \cdots & x_{k+i+j-2} \\ u_{k+i} & \cdots & u_{k+i+j-2} \end{bmatrix}$$

辨识。

上述结果构成了两步辨识方法的核心。

第一步, 由 \mathbf{H}_1 和 \mathbf{H}_2 张成的行空间的交集可以通过级联矩阵 $\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix}$ 的奇异值分解求出, 即

$$\mathbf{H} = \mathbf{U}_H \boldsymbol{\Sigma}_H \mathbf{V}_H^T = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{V}_H^T$$

其中

$$\dim(\mathbf{U}_{11}) = (mi + li) \times (2mi + n)$$

$$\dim(\mathbf{U}_{12}) = (mi + li) \times (2li - n)$$

$$\dim(\mathbf{U}_{21}) = (mi + li) \times (2mi + n)$$

$$\dim(\mathbf{U}_{22}) = (mi + li) \times (2li - n)$$

$$\dim(\boldsymbol{\Sigma}_{11}) = (2mi + n) \times (2mi + n)$$

根据

$$\mathbf{U}_{12}^T \mathbf{H}_1 = -\mathbf{U}_{22}^T \mathbf{H}_2$$

可以得出结论: $\mathbf{U}_{12}^T \mathbf{H}_1$ 等于所需要的交集。然而, $\mathbf{U}_{12}^T \mathbf{H}_1$ 包含有 $2li - n$ 个行向量, 其中, 只有 n 个 (交集的维数) 是线性无关的。因此, 需要选择这些行向量的适当组合。

第二步, 由于 \mathbf{U}_{12} 和 \mathbf{U}_{22} 构成一个正交矩阵, 所以它们可以奇异值分解为^[184]

$$\begin{aligned} \mathbf{U}_{12} &= [\mathbf{U}_{12}^{(1)}, \mathbf{U}_{12}^{(2)}, \mathbf{U}_{12}^{(3)}] \begin{bmatrix} \mathbf{I}_{(li-n) \times (li-n)} & & \\ & \mathbf{C}_{n \times n} & \\ & & \mathbf{O}_{(li-n) \times (li-n)} \end{bmatrix} \mathbf{V}^T \\ \mathbf{U}_{22} &= [\mathbf{U}_{22}^{(1)}, \mathbf{U}_{22}^{(2)}, \mathbf{U}_{22}^{(3)}] \begin{bmatrix} \mathbf{O}_{(li-n) \times (li-n)} & & \\ & \mathbf{S}_{n \times n} & \\ & & \mathbf{I}_{(li-n) \times (li-n)} \end{bmatrix} \mathbf{V}^T \end{aligned}$$

其中

$$\mathbf{C} = \text{diag}(c_1, \dots, c_n)$$

$$\mathbf{S} = \text{diag}(s_1, \dots, s_n)$$

$$\mathbf{I}_{n \times n} = \mathbf{C}^2 + \mathbf{S}^2$$

于是, $U_{12}^{(1)}$ 构成 H_1 的 $(li-n)$ 维正交补。显然, 只有 $U_{12}^{(2)}$ 组成交集计算的有用组合, 因而可以取

$$x = (U_{12}^{(2)})^H H_1$$

上述两步方法是一种双重奇异值分解辨识算法, 即所需要的交集先后简化为 H 和 U_{12} 两个矩阵的奇异值分解。

6.8.3 阶数确定

考虑线性移不变的平稳随机自回归-滑动平均 (ARMA) 过程 $\{x(n)\}$:

$$x(n) + \sum_{i=1}^p a(i)x(n-i) = \sum_{j=0}^q b(j)e(n-j), \quad e(n) \sim \text{IID}(0, \sigma_e^2) \quad (6.8.14)$$

式中, $e(n) \sim \text{IID}(0, \sigma_e^2)$ 表示 $\{e(n)\}$ 是独立同分布的随机过程, 其均值为 0, 方差等于 σ_e^2 。令 $R(\tau) = E\{x(n)x^*(n+\tau)\}$ 是 $x(n)$ 的自相关函数。我们的目的是只根据观测数据确定 ARMA 模型的自回归 (AR) 阶次 p 和滑动平均 (MA) 阶次 q 。

确定 ARMA 模型阶次的方法已提出不少。归纳起来, 可以分为两大类: 信息量准则法和线性代数法。

信息量准则法有著名的 AIC 准则, 最终预测误差 (FPE) 准则, 最小描述长度 (MDL) 准则, 准则自回归传递 (CAT) 准则和 BIC (AIC 的改进) 准则等, 有关这些准则的详细综合介绍可在文献 [524, pp.90~91] 中找到。

典型的线性代数定阶法有行列式检验算法^[85]、Gram-Schmidt 正交法^[76]和奇异值分解法^{[70],[506],[507]}。

- (1) 行列式检验法检验不同维数的行列式的值, 以行列式等于零时的最小维数作为阶数的估计。
- (2) Gram-Schmidt 正交法则检验与 ARMA 模型服从的法方程对应的矩阵的列之间的线性无关性, 以线性无关的列数作为阶数的估计。
- (3) 奇异值分解法确定与 ARMA 模型服从的法方程对应的矩阵的有效秩, 并用它作阶数的估计。

从数值稳定性看, 前两种方法都明显比奇异值分解方法差。

对于式 (6.8.14) 的 ARMA 模型, 下列著名的修正 Yule-Walker (MYW) 方程成立 (参见文献 [524]):

$$\sum_{i=1}^p a(i)R(\tau-i) = 0, \quad i > q \quad (6.8.15)$$

利用 MYW 方程, Cadzow 证明了以下结果^[70]。

命题 6.8.1 记 $t \times (p_e + 1)$ 矩阵

$$R_e = \begin{bmatrix} R(q_e + 1) & R(q_e) & \cdots & R(q_e - p_e + 1) \\ R(q_e + 2) & R(q_e + 1) & \cdots & R(q_e - p_e + 2) \\ \vdots & \vdots & & \vdots \\ R(q_e + t) & R(q_e + t - 1) & \cdots & R(q_e - p_e + t) \end{bmatrix} \quad (6.8.16)$$

其中, $t \geq p_e$. 若 $p_e \geq p, q_e \geq q$, 且 $q_e - p_e \geq q - p$, 则 $\text{rank}(\mathbf{R}_e) = p$.

这一结果给出了一种确定 ARMA 模型式 (6.8.14) 的 AR 阶数 p 的方法. 理论自相关函数 $R(\tau)$ 用其样本估计值 $\hat{R}(\tau) = \frac{1}{N} \sum_{n=1}^{N-\tau} x(n)x(n+\tau)$ 代替后, 样本自相关矩阵 $\hat{\mathbf{R}}_e$ 的有效秩即给出 AR 阶数 p 的估计, 而 $\hat{\mathbf{R}}_e$ 的有效秩可以利用 $\hat{\mathbf{R}}_e$ 的奇异值分解来确定. 顺便指出, 选择满足 $p_e \geq p, q_e \geq q$ 和 $q_e - p_e \geq q - p$ 的扩展阶 p_e 和 q_e 并不困难, 因为不失一般性, 通常可以假定 $q \leq p$, 此时可以直接选择 $p_e = q_e \gg p$.

在确定出 AR 阶数 p 后, MA 阶数 q 也可以利用 SVD 确定. 利用 SVD 确定 ARMA 模型式 (6.8.14) 的 MA 阶数有两种方法, 由 Zhang X. D. 与 Zhang Y. S. [506],[507] 提出. 两种方法的主要区别在于第一种方法只需要使用 AR 阶次 p , 而无需使用 ARMA 模型的任何其他参数; 第二种方法则要用到 AR 阶次 p 和 AR 参数 $a(i), i = 1, 2, \dots, p$.

下面介绍这两种方法.

1. 搜索法

这一方法 [506] 利用了这样一个事实: MA 阶数 q 的信息隐含在另外一个自相关矩阵内, 该矩阵可以看作是 \mathbf{R}_e 的增广矩阵. 更严格地, 这一事实可叙述为下面的命题.

命题 6.8.2 [506] 令 \mathbf{R}_1 是一个由下式定义的 $(p+1) \times (p+1)$ 维矩阵:

$$\mathbf{R}_1 = \begin{bmatrix} R(q) & R(q-1) & \cdots & R(q-p) \\ R(q+1) & R(q) & \cdots & R(q-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ R(q+p) & R(q+p-1) & \cdots & R(q) \end{bmatrix} \quad (6.8.17)$$

若 $a_p \neq 0$, 则 $\text{rank}(\mathbf{R}_1) = p+1$.

上述结果应用起来是不方便的, 因为它要用到待确定的 MA 阶次 q . 幸运的是, 我们很容易把它推广到具有扩展阶 $q_e > q$ 的矩阵去.

命题 6.8.3 [506] 假定 p 已经确定出, 且 $q_e \geq q$, 并令 \mathbf{R}_{1e} 是一个 $(p+1) \times (p+1)$ 维矩阵, 且

$$\mathbf{R}_{1e} = \begin{bmatrix} R(q_e) & R(q_e-1) & \cdots & R(q_e-p) \\ R(q_e+1) & R(q_e) & \cdots & R(q_e-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ R(q_e+p) & R(q_e+p-1) & \cdots & R(q_e) \end{bmatrix} \quad (6.8.18)$$

则当 $q_e > q$ 时, $\text{rank}(\mathbf{R}_{1e}) = p$, 而且仅当 $q_e = q$ 时才有 $\text{rank}(\mathbf{R}_{1e}) = p+1$.

显然, 真实的 MA 阶数 q 隐含在 \mathbf{R}_{1e} 内. 理论上, 阶数 q 可以这样来确定: 从某个阶数 $Q = q_e > q$ 开始, 用 SVD 检验 \mathbf{R}_{1e} 的秩, 然后依次减小一阶 ($Q \leftarrow Q-1$) 进行新的检验. 秩从 p 变为 $p+1$ 的第一个转折点发生在 $Q = q$ 处. 然而, 在实际应用中, 由于只能使用样本自相关估计值, 所以阶数从 p 变为 $p+1$ 的转折点往往不是很明显. 为了发展一种 MA 定阶的实际算法, 可以使用“超定的”矩阵 \mathbf{R}_{2e} , 其元素 $\mathbf{R}_{2e}[i, j] = \hat{R}(q_e + i - j), i = 1, 2, \dots, t, j = 1, 2, \dots, p_e + 1 (t \gg p_e)$, 这意味着使用更多的

样本自相关函数。容易验证:

$$\text{rank}(\mathbf{R}_{2e}) = \text{rank}(\mathbf{R}_{1e})$$

因为 \mathbf{R}_{2e} 包含了整个 \mathbf{R}_{1e} , 另根据修正 Yule-Walker 方程 (6.8.15), 任意第 $k(\geq p+2)$ 行 (或列) 与其上面 p 行 (或左边 p 列) 线性相关。

下面是 MA 定阶的搜索方法。

步骤 1 用 SVD 确定 AR 阶次 p , 并取 $Q = q_e > q$ 。

步骤 2 $Q \leftarrow Q-1$, 并利用样本自相关函数构造样本自相关矩阵 $\hat{\mathbf{R}}_{2e}$, 计算其 SVD。

步骤 3 如果第 $(p+1)$ 个奇异值与上次计算结果相比, 有一个明显的转折, 则选取 $q = Q$; 否则返回步骤 2, 并重复以上步骤, 直到 q 被选出。

定阶算法的关键是在步骤 3 中选择第 $(p+1)$ 个奇异值的转折点。作为转折检验法, 可以采用检验该奇异值前后的相对变化率

$$\alpha = \frac{\sigma_{p+1}^Q - \sigma_{p+1}^{Q+1}}{\sigma_{p+1}^{Q+1}}$$

是否大于某个预先设定的阈值的方法。若 α 大于该阈值, 则接受此 Q 值为转折点。

2. 拟合残差法

这种 MA 定阶方法需要使用 AR 阶次 p 和 AR 参数 $a(i), i = 1, 2, \dots, p$ (有关 AR 参数估计的 SVD-TLS 方法, 将在下一章介绍)。构造拟合残差

$$f(\tau) = \sum_{i=0}^p a(i)R(\tau-i) \quad (6.8.19)$$

根据 MYW 方程易知

$$f(\tau) \begin{cases} \neq 0, & \tau = 0, 1, \dots, q \\ = 0, & \tau = q+1, q+2, \dots \end{cases} \quad (6.8.20)$$

原理上, MA 阶次 q 是满足 $f(\tau) \equiv 0$ 的最小整数 τ 。然而, 在实际应用中, 由于根据 AR 参数估计值和样本自相关函数计算的 $\hat{f}(\tau)$ 值存在比较大的误差, 所以直接根据 $\hat{f}(\tau) \approx 0$ 确定 q 是一种数值鲁棒性差的定阶方法。文献 [507] 提出利用拟合残差构造一特殊矩阵, 该矩阵的秩直接给出 q , 因此无需计算多个自相关矩阵的 SVD。

方法是简单的。容易看出三角矩阵

$$\mathbf{R}_q = \begin{bmatrix} f(q) & f(q-1) & \cdots & f(0) \\ 0 & f(q) & \cdots & f(1) \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(q) \end{bmatrix} \quad (6.8.21)$$

是满秩的, 其秩等于 $q+1$, 因为对角线元素 $f(q) \neq 0$ 。与搜索法类似, 也可以构造一个

$(q_e + 1) \times (q_e + 1)$ 扩展阶矩阵

$$R_{q_e} = \begin{bmatrix} f(q_e) & f(q_e - 1) & \cdots & f(0) \\ 0 & f(q_e) & \cdots & f(1) \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(q_e) \end{bmatrix} \quad (6.8.22)$$

它的元素不再含有未知的 q , 但矩阵的秩却仍然等于 $q + 1$ 。

利用式 (6.8.20) 容易证明, $\text{rank}(R_{q_e}) = \text{rank}(R_q) = q + 1$ 。因此, 利用 SVD 确定出相应的样本残差矩阵 \hat{R}_{q_e} 的有效秩后, 即可直接得到 MA 阶次 q 。有趣的是, 确定矩阵 \hat{R}_{q_e} 的秩显然等价于检验对角线元素乘积 (PODE) 是否等于零, 即

$$f^{\tau+1}(\tau) \begin{cases} \neq 0, & \tau = q \\ \equiv 0, & \tau = q + 1, q + 2, \dots \end{cases}$$

具体说来, PODE 试验就是选择使 $f^{\tau+1}(\tau) \equiv 0$ 的最小整数作为 MA 阶次 q 的估计。PODE 试验明显地比直接检验式 (6.8.20) 具有更好的数值鲁棒性。SVD 方法和 PODE 实验还可综合起来使用, 这样可以进一步提高 MA 阶数确定的数值鲁棒性, 详见文献 [507]。

顺便指出, 上述使用自相关函数确定 AR 和 MA 阶次的方法仅适用于因果最小相位的 ARMA 过程。对于非最小相位系统, 应该使用高阶累积量定阶和参数估计。基于高阶累积量确定 AR 和 MA 阶数的 SVD 方法分别由文献 [171] 和 [507] 提出。对这两种方法感兴趣的读者也可参考文献 [521]。文献 [507] 和 [521] 还介绍了使用 SVD 定阶方法进行高斯和非高斯过程识别, 以及在 SVD 定阶方法在 AR, MA 和 ARMA 模型辨别中的应用。

6.8.4 系统的可控性

考查线性系统

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \quad \mathbf{x}(0) = \mathbf{x}_0$$

其中, $\mathbf{A} \in R^{n \times n}$, $\mathbf{b} \in R^n$, $\mathbf{x}_0 \in R^n$ 。我们的问题是求控制函数 $u(t)$, 使 $\mathbf{x}(T) = \mathbf{x}_T$ (其中, $\mathbf{x}_T \in R^n$) 是某个希望的“状态”, 即我们希望系统在 $T > 0$ 时处于某种状态。显然, 这有可能是不能实现的。例如, 若 $\mathbf{b} = \mathbf{0}$, 则将无法像希望的那样对 \mathbf{x} 向量加以控制。更一般地, 如果上述问题有解, 则 \mathbf{b} 在可以用 \mathbf{A} 的特征系统定义的方向上不得是秩亏缺的。有多种方法可以描述一个系统 S 的“可控性”, 例如

(1) 系统 S 是可控的, 当且仅当 $\mathbf{W}_1 = [\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \dots, \mathbf{A}^{n-1}\mathbf{b}]$ 非奇异。

(2) 系统 S 是可控的, 当且仅当 $\mathbf{W}_2 = \int_0^T \mathbf{e}^{\mathbf{A}t} \mathbf{b} \mathbf{b}^T \mathbf{e}^{\mathbf{A}^T t} dt$ 非奇异。

它们是“0-1”或“是非”描述方法。直觉告诉我们, 如果 \mathbf{W}_1 或 \mathbf{W}_2 接近奇异, 那么 S 必然“难于”控制。这种情况很容易用奇异值进行分析。可以用 \mathbf{W}_1 或 \mathbf{W}_2 的最小奇异值作为系统 S 非可控的测度。在鲁棒系统的设计中, 度量系统非可控的测度起着重要的作用。关于数值线性代数在控制理论中的作用的精彩解说可在文献 [350], [516] 中找到。

6.8.5 图像压缩

奇异值分解在图像处理中有着重要应用。假定一幅图像有 $n \times n$ 个像素, 如果将这 n^2 个数据一起传送, 往往会显得数据量太大。因此, 我们希望能够改为传送另外一些比较少的数据, 并且在接收端还能够利用这些传送的数据重构原图像。

不妨用 $n \times n$ 矩阵 A 表示要传送的原 $n \times n$ 个像素。假定对矩阵 A 进行奇异值分解, 便得到 $A = U \Sigma V^T$, 其中, 奇异值按照从大到小的顺序排列。如果从中选择 k 个大奇异值以及与这些奇异值对应的左和右奇异向量逼近原图像, 便可以共使用 $k(2n+1)$ 个数值代替原来的 $n \times n$ 个图像数据。这 $k(2n+1)$ 个被选择的新数据是矩阵 A 的前 k 个奇异值、 $n \times n$ 左奇异向量矩阵 U 的前 k 列和 $n \times n$ 右奇异向量矩阵 V 的前 k 列的元素。

比率

$$\rho = \frac{n^2}{k(2n+1)} \quad (6.8.23)$$

称为图像的压缩比。显然, 被选择的大奇异值的个数 k 应该满足条件 $k(2n+1) < n^2$ 即 $k < \frac{n^2}{2n+1}$ 。因此, 我们在传送图像的过程中, 就无需传送 $n \times n$ 个原始数据, 而只需要传送 $k(2n+1)$ 个有关奇异值和奇异向量的数据即可。在接收端, 在接收到奇异值 $\sigma_1, \sigma_2, \dots, \sigma_k$ 以及左奇异向量 u_1, u_2, \dots, u_k 和右奇异向量 v_1, v_2, \dots, v_k 后, 即可通过截尾的奇异值分解公式

$$\hat{A} = \sum_{i=1}^k \sigma_i u_i v_i^T \quad (6.8.24)$$

重构出原图像。

一个容易理解的事实是: 若 k 值偏小, 即压缩比 ρ 偏大, 则重构的图像的质量有可能不能令人满意。反之, 过大的 k 值又会导致压缩比过小, 降低图像压缩和传送的效率。因此, 需要根据不同种类的图像, 选择合适的压缩比, 以兼顾图像传送效率和重构质量。

6.9 广义奇异值分解的应用

仍然考虑式 (6.8.11) 的系统的辨识。但是, 不同的是, 这里的输入/输出被有色噪声所污染。假定级联矩阵

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}$$

的列存在零均值的独立同分布误差, 它们之间的误差协方差矩阵为 Δ (包含一个比例因子), 其中

$$\Delta = R_{\Delta} R_{\Delta}^H$$

是矩阵 Δ 的 Cholesky 分解, 分解因子 R_{Δ} 是下三角矩阵, 常称为矩阵 Δ 的平方根。

容易验证, 变换后的矩阵 $R_{\Delta}^{-1} H$ 的列的误差协方差矩阵等于单位矩阵 (包含一比例因子)。因此, 系统辨识的方法之一就是利用 $R_{\Delta}^{-1} H$ 的奇异值分解代替 H 的奇异值分解求

出矩阵 A , B , C 和 D 。然而, 若 R_{Δ} 奇异或病态, 则 R_{Δ}^{-1} 将不能够直接计算。回顾广义奇异值分解知, 应该用矩阵束 (H, R_{Δ}) 的广义奇异值分解取代 $R_{\Delta}^{-1}H$ 的普通奇异值分解。于是, 很容易将 6.8.2 小节的双重奇异值分解系统辨识方法推广成一种双重广义奇异值分解的系统辨识方法^[424]。

令矩阵束 (H, R_{Δ}) 的广义奇异值分解为

$$\begin{aligned} X^T H U &= \Sigma_H \\ X^T R_{\Delta} V &= \Sigma_{R_{\Delta}} \end{aligned}$$

式中

$$\begin{aligned} \Sigma_H &= \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_{2li+2mi}) \\ \Sigma_{R_{\Delta}} &= \text{diag}(\beta_1, \beta_2, \dots, \beta_{2li+2mi}) \\ \frac{\alpha_1}{\beta_1} &> \frac{\alpha_2}{\beta_2} > \dots > \frac{\alpha_{2li+2mi}}{\beta_{2li+2mi}} \end{aligned}$$

现在的情况与 6.8.2 小节的情况非常类似。在 6.8.2 节, H_1 和 H_2 的行空间的交集是利用级联矩阵 $H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}$ 的最小有向信号能量 (即 $\begin{bmatrix} U_{12} \\ U_{22} \end{bmatrix}$, 且 $U_{12}^T H_1 = -U_{22}^T H_2$) 计算的。因此, 在现在的情况下, 可以利用 $\begin{bmatrix} X_{12} \\ X_{22} \end{bmatrix}$ (其定义在后面给出) 计算 H_1 和 H_2 的行空间的交集。

先考虑无噪声的情况 (误差的协方差与 Δ 成正比, 但比例因子为零), 然后再说明推导的结果对于非零的误差分布仍然成立。

如果数据是无噪声的, 则根据上述广义奇异值分解知

$$H_1 = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = X^{-T} \Sigma_H U^T = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{11} & O \\ O & O \end{bmatrix} U^T$$

式中, 各个分块矩阵的维数分别为

$$\begin{aligned} \dim(X_{11}) &= (mi + li) \times (2mi + n) \\ \dim(X_{12}) &= (mi + li) \times (2li - n) \\ \dim(X_{21}) &= (mi + li) \times (2mi + n) \\ \dim(X_{22}) &= (mi + li) \times (2li - n) \\ \dim(\Sigma_{11}) &= (2mi + n) \times (2mi + n) \end{aligned}$$

由

$$X_{12}^T H_1 = -X_{22}^T H_2$$

可知, $X_{12}^T H_1$ 的行空间等于欲求的交集。由于 $X_{12}^T H_1$ 包含 $2li - n$ 个行向量, 而其中只有 n 个 (交集的维数) 行向量线性无关, 所以我们需要从 $2li - n$ 个行向量中选出 n 个合

适的向量来。利用广义奇异值分解, 容易证明:

$$\begin{aligned} \mathbf{X}_{12} &= [\mathbf{X}_{12}^{(1)}, \mathbf{X}_{12}^{(2)}, \mathbf{X}_{12}^{(3)}] \begin{bmatrix} \mathbf{I}_{(li-n) \times (li-n)} & \mathbf{C}_{n \times n} & \mathbf{O}_{(li-n) \times (li-n)} \end{bmatrix} \mathbf{Q}^T \\ \mathbf{X}_{22} &= [\mathbf{X}_{22}^{(1)}, \mathbf{X}_{22}^{(2)}, \mathbf{X}_{22}^{(3)}] \begin{bmatrix} \mathbf{O}_{(li-n) \times (li-n)} & \mathbf{S}_{n \times n} & \mathbf{I}_{(li-n) \times (li-n)} \end{bmatrix} \mathbf{Q}^T \end{aligned}$$

其中, \mathbf{Q} 为正交矩阵, 并且

$$\mathbf{C} = \text{diag}(c_1, c_2, \dots, c_n)$$

$$\mathbf{S} = \text{diag}(s_1, s_2, \dots, s_n)$$

$$\mathbf{I}_{n \times n} = \mathbf{C}^2 + \mathbf{S}^2$$

显然, 只有 $\mathbf{X}_{12}^{(2)}$ 提供计算交集的有用向量组合, 故可以取

$$\mathbf{x} = \mathbf{X}_{12}^{(2)} \mathbf{H}_1$$

现在, 系统矩阵可以通过下述简化的矩阵方程 (忽略共同的正交因子 \mathbf{U}) 计算:

$$\begin{aligned} & \begin{bmatrix} (\mathbf{X}_{12}^{(2)})^T \mathbf{X}^{-T}(m+l+1:(i+1)(m+l), 1:2mi+n) \boldsymbol{\Sigma}_{11} \\ \mathbf{X}^{-T}(mi+li+m+1:(m+l)(i+1), 1:2mi+n) \boldsymbol{\Sigma}_{11} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} (\mathbf{X}_{12}^{(2)})^T \mathbf{X}^{-T}(1:mi+li, 1:2mi+n) \boldsymbol{\Sigma}_{11} \\ \mathbf{X}^{-T}(mi+li+1:mi+li+m, 1:2mi+n) \boldsymbol{\Sigma}_{11} \end{bmatrix} \end{aligned}$$

式中, $\mathbf{X}^{-T}(r:s, v:w)$ 代表 \mathbf{X}^{-T} 的子矩阵, 它具有 \mathbf{X}^{-T} 的第 $r, r+1, \dots, s$ 行和第 $v, v+1, \dots, w$ 列。

如果在 \mathbf{H}_1 和 \mathbf{H}_2 内的测量数据存在某种不精确度, 则所有的广义奇异值都是非零的。此时, 需要将最小的广义奇异值置零。应该指出的是, 当 \mathbf{H} 的列数趋于无穷, 且 \mathbf{H} 的各列是具有零均值和相同的误差协方差矩阵 $\boldsymbol{\Delta}$ (包含一比例因子) 的独立同分布误差时, 广义奇异值分解中的矩阵 \mathbf{X} 可以一致计算。由于系统矩阵只根据 \mathbf{X} 计算, 所以系统模型的估计是一致的。

信号处理中著名的 MUSIC (多重信号分类) 方法等价于广义奇异值分解问题。另一著名的 ESPRIT (借助旋转不变技术估计信号参数) 方法是广义特征值分解的巧妙应用。我们将在第 8 章详细介绍这两种应用。

本章小结

本章对矩阵的奇异值分析展开了专题讨论。首先分析了单个矩阵的 (普通) 奇异值分解 (OSVD)、奇异值的性质以及奇异值分解的数值计算。然后, 以两个或者多个矩阵为对象, 介绍了各种推广的奇异值分解:

- (1) 两个矩阵乘积 $A^T B$ 的奇异值分解, 简称乘积奇异值分解 (PSVD);
- (2) 两个矩阵组成的矩阵对 (A, B) 的广义奇异值分解 (GSVD);
- (3) 矩阵三元组 (A, B, C) 的约束奇异值分解 (RSVD);
- (4) 矩阵相对于某种分块结构的结构奇异值。

普通奇异值分解、乘积奇异值分解和广义奇异值分解都是约束奇异值分解的特例。

本章最后两节分别介绍了奇异值分解在静态系统建模、系统辨识、阶数确定、系统的可控性、图像压缩中的应用, 以及广义奇异值分解在系统辨识中的应用。

习 题

6.1 已知矩阵

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

通过计算 AA^T 和 $A^T A$ 的特征值和特征向量, 求矩阵 A 的奇异值分解。

6.2 计算矩阵

$$A = \begin{bmatrix} 1 & -1 \\ 3 & -3 \\ -3 & 3 \end{bmatrix}$$

的奇异值分解。

6.3 计算矩阵

$$A = \begin{bmatrix} 3 & 4 & 5 \\ 2 & 1 & 7 \end{bmatrix}$$

的奇异值分解。

6.4 已知矩阵

$$A = \begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & 9 & -25 \end{bmatrix}$$

求 A 的奇异值以及最小奇异值 σ_1 相对应的左、右奇异向量。

6.5 令 $A = xp^H + yq^H$, 其中, $x \perp y$ 和 $p \perp q$ 。求矩阵 A 的 Frobenius 范数 $\|A\|_F$ 。(提示: 计算 $A^H A$, 并求 A 的奇异值。)

6.6 已知 $A = U \Sigma V^H$ 是矩阵 A 的奇异值分解, 矩阵 A^H 的奇异值与 A 的奇异值有何关系?

6.7 证明: 若 A 为正方矩阵, 则 $|\det(A)|$ 等于 A 的奇异值之积。

6.8 假定 A 为可逆矩阵, 求 A^{-1} 的奇异值分解。

6.9 证明: 若 A 为 $n \times n$ 正定矩阵, 则 A 的奇异值与 A 的特征值相同。

6.10 令 A 为 $m \times n$ 矩阵, 且 P 为 $m \times m$ 正交矩阵。证明 PA 与 A 的奇异值相同。矩阵 PA 与 A 的左、右奇异向量有何关系?

6.11 令 A 是一个 $m \times n$ 矩阵, 并且 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是矩阵 $A^T A$ 的特征值, 相对应的特征向量为 u_1, u_2, \dots, u_n 。证明 A 的奇异值 σ_i 等于范数 $\|Au_i\|$, 即

$$\sigma_i = \|Au_i\|, \quad i = 1, 2, \dots, n$$

6.12 令 $\lambda_1, \lambda_2, \dots, \lambda_n$ 和 u_1, u_2, \dots, u_n 分别是矩阵 $A^T A$ 的特征值和特征向量。假定矩阵 A 有 r 个非零的奇异值, 证明 $\{Au_1, Au_2, \dots, Au_r\}$ 是列空间 $\text{Col}(A)$ 的一组正交基, 并且 $\text{rank}(A) = r$ 。

6.13 令 $B, C \in R^{m \times n}$, 求复矩阵 $A = B + jC$ 与实分块矩阵 $\begin{bmatrix} B & -C \\ C & B \end{bmatrix}$ 的奇异值和奇异向量之间的关系。

6.14 用矩阵 $A \in R^{m \times n}$ ($m \geq n$) 的奇异向量表示 $\begin{bmatrix} O & A^T \\ A & O \end{bmatrix}$ 的特征向量。

6.15 利用 MATLAB 函数 $[U, S, V] = \text{svd}(X)$ 求解方程 $Ax = b$, 其中

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & 3 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 2 \end{bmatrix}$$

6.16 假定计算机仿真的观测数据为

$$x(n) = \sqrt{20} \sin(2\pi 0.2n) + \sqrt{2} \sin(2\pi 0.215n) + w(n)$$

产生, 其中, $w(n)$ 是一高斯白噪声, 其均值为 0, 方差为 1, 并取 $n = 1, 2, \dots, 128$ 。试针对 10 次独立的仿真实验数据, 分别确定自相关矩阵

$$R = \begin{bmatrix} r(0) & r(-1) & \cdots & r(-2p) \\ r(1) & r(0) & \cdots & r(-2p+1) \\ \vdots & \vdots & \ddots & \vdots \\ r(M) & r(M-1) & \cdots & r(M-2p) \end{bmatrix}$$

的有效秩。式中, $r(k) = \frac{1}{128} \sum_{i=1}^{128-k} x(i)x(i+k)$ 表示观测信号的样本自相关函数 (未知的观测数据皆令其等于 0), 并取 $M = 50, p = 10$ 。

6.17 ^[184] 使用奇异值分解证明: 若 $A \in R^{m \times n}$ ($m \geq n$), 则存在 $Q \in R^{m \times n}$ 和 $P \in R^{n \times n}$, 使得 $A = QP$, 其中, $Q^T Q = I_n$, 并且 P 是对称的和非负定的。这一分解有时称为极分解 (polar decomposition), 因为它与复数分解 $z = |z|e^{j\arg(z)}$ 类似。