

一类矩阵的逆奇异值问题

吴春红*, 卢琳璋**

(厦门大学数学科学学院, 福建 厦门 361005)

摘要: 讨论了一类矩阵的逆奇异值问题. 给定非负实数 $\sigma_1, \sigma_2, \dots, \sigma_n$, 两非零实向量 $x = (x_1, x_2, \dots, x_m)^T, y = (y_1, y_2, \dots, y_n)^T$, 求 $m \times n$ 阶实矩阵 A , 使得 $\sigma_1, \sigma_2, \dots, \sigma_n$ 为 A 的奇异值, 并且 x, y 分别为 A 的左右奇异向量. 基于 Householder 变换和矩阵秩 1 的修正方法得到了问题的算法, 而且算法比较经济且易于并行, 同时给出了相应的数值例子.

关键词: 反问题; 奇异值; Householder 变换; 秩 1 修正

中图分类号: O 241.6

文献标识码: A

文章编号: 0438-0479(2010)03-0308-04

逆奇异值问题为矩阵逆特征值问题的推广, 通常考虑的是由给定的奇异值来构造结构矩阵. 逆奇异值问题在确定物体的质量分布、轨道力学、灌溉理论、断层扫描、电路理论等领域有重要的应用^[1-3]. 若要构造的矩阵为对称阵, 则逆奇异值问题就转化为逆特征值问题, 其应用参见文献[4-8], 因此逆奇异值问题要求构造的矩阵不为对称阵, 但可以是长方阵, 这使得逆奇异值问题更为复杂. 设矩阵 A 的奇异值为 $\sigma_i (i = 1, 2, \dots, n)$, 则对称阵 $\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$ 的特征值为 $\alpha_i (i = 1, 2, \dots, n)$, 即可将逆奇异值问题转化为一类特殊的逆特征值问题, 因此常将逆奇异值问题归类为逆特征值问题.

本文中 e_j 在有意义的前提下为相应阶数单位阵的第 j 列; $\mathbf{R}^{m \times n}$ 表示 $m \times n$ 阶实矩阵.

Chu 在 1992 年讨论了如下逆奇异值问题^[9]:

给定矩阵 $\{A_i\}_{i=0}^n \in \mathbf{R}^{m \times n} (m \geq n)$ 及非负实数 $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$, 求 $c = (c_1, c_2, \dots, c_n)^T \in \mathbf{R}^n$ 使得

$$A(c) = A_0 + \sum_{i=1}^n c_i A_i$$

的奇异值为 $\alpha_1, \alpha_2, \dots, \alpha_n$.

Chu 提出了两种数值方法来求解上述逆奇异值问题: 其一是用投影梯度法来求解目标函数的常微分方程的; 另一种将 Friedland^[10] 提出的迭代算法推广, 提出了 Newton-type 方法. Bai 提出了用非精确的 Newton-type 方法来求解上述问题^[11], 并证明了此算法超

线性收敛. 最近, 一些学者提出了不同的逆奇异值问题^[12]. 类似于极点配置问题, Chu 提出了如下逆奇异值问题^[13]:

给定矩阵 $\{A_i\}_{i=0}^n \in \mathbf{R}^{m \times n} (m \geq n)$, 且 $n \geq l > 0$, 及非负实数 $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$, 求矩阵 $F \in \mathbf{R}^{m \times n}$, 使得 $\text{rank}(F) \leq l$, 且 $A + F$ 的奇异值为 $\alpha_1, \alpha_2, \dots, \alpha_n$. 即用低秩阵 F 来修正 A 的奇异值.

Zhu 提出用 Householder 变换和秩 1 修正来求解对称阵的逆特征值问题^[14]. 本文提出用 Householder 变换和秩 1 修正来求解如下逆奇异值问题:

问题 I 给定非负实数 $\alpha_1, \alpha_2, \dots, \alpha_n$, 两非零向量 $x \in \mathbf{R}^m, y \in \mathbf{R}^n$, 求矩阵 $A \in \mathbf{R}^{m \times n}$ 使得 $\sigma_1, \sigma_2, \dots, \sigma_n$ 为 A 的奇异值, 且 x, y 分别为 A 的左右奇异向量.

1 问题的解与算法

假定 $m \geq n$, 矩阵 $A \in \mathbf{R}^{m \times n}$, 则存在正交阵 $U = [u_1, u_2, \dots, u_m] \in \mathbf{R}^{m \times m}$ 和正交阵 $V = [v_1, v_2, \dots, v_n] \in \mathbf{R}^{n \times n}$, 使得

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^T, \quad (2)$$

其中 $\Sigma = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbf{R}^{n \times n}$, 且 $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n \geq 0$. 式(2)称为矩阵 A 的奇异值分解(SVD)^[15-16]; 非负实数 $\alpha_i (i = 1, 2, \dots, n)$ 称为矩阵 A 的奇异值, u_i 为对应于奇异值 α_i 的左奇异向量, v_i 为对应于奇异值 α_i 的右奇异向量. 实际上可验证 $\alpha_i^2 (i = 1, 2, \dots, n)$ 为对称阵 $A^T A \in \mathbf{R}^{n \times n}$ 的特征值, 右奇异向量 v_i 为对应于 α_i 的单位特征向量. 对称阵 $AA^T \in \mathbf{R}^{m \times m}$ 的特征值为 α_i^2 和 0 ($m - n$ 个), 左奇异向量 u_i 为对应于 α_i^2 的单位特征向量.

引理 1^[17] 若 $\lambda_1, \lambda_2, \dots, \lambda_n$ 为 n 阶矩阵 A 的特征

收稿日期: 2009-07-22

基金项目: 国家自然科学基金(10531080); 厦门理工学院科学技术研究项目(YKJ09016R)

* 现工作单位: 厦门理工学院

** 通讯作者: lzlu@xmu.edu.cn

值, u 为对应于特征值 λ 的特征向量, $v \in \mathbf{R}^n$ 为任意向量, 则 $A + uv^T$ 为矩阵 A 的秩 1 修正, 其特征值为

$$\lambda, \dots, \lambda_{-1}, \lambda + v^T u, \lambda_{+1}, \dots, \lambda.$$

定理 1 假定矩阵 $A \in \mathbf{R}^{m \times n}$ ($m \geq n$), 其奇异值分解为式(2), 设 $\rho > 0$, 则 $A + \rho u_i v_i^T$ ($i = 1, 2, \dots, n$) 的奇异值为 $\alpha_i, \alpha_2, \dots, \alpha_i + \rho, \alpha_{i+1}, \dots, \alpha_n$, 且对应地左右奇异向量不变.

证明 令 $B = A + \rho u_i v_i^T$, 由式(2), 得

$$B^T B = (A + \rho u_i v_i^T)^T (A + \rho u_i v_i^T) =$$

$$A^T A + \rho A^T u_i v_i^T + \rho u_i^T A + \rho^2 v_i v_i^T =$$

$$A^T A + \rho V \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix}^T U^T u_i v_i^T +$$

$$\rho v_i u_i^T U \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix} V^T + \rho^2 v_i v_i^T =$$

$$A^T A + \rho V \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix}^T e_i v_i^T +$$

$$\rho v_i e_i^T \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix} V^T + \rho^2 v_i v_i^T =$$

$$A^T A + \rho \alpha_i v_i v_i^T + \rho \alpha_i v_i v_i^T + \rho^2 v_i v_i^T =$$

$$A^T A + 2\rho \alpha_i v_i v_i^T + \rho^2 v_i v_i^T,$$

因 α_i^2 ($i = 1, 2, \dots, n$) 为矩阵 $A^T A$ 的特征值, 且对应的特征向量为 v_i ($i = 1, 2, \dots, n$). 由引理 1, 所以矩阵 $B^T B$ 的特征值为 $\alpha_i^2, \dots, \alpha_i^2 + 2\rho \alpha_i + \rho^2, \alpha_{i+1}^2, \dots, \alpha_n^2$, 因此 B 的奇异值为 $\alpha_i, \alpha_2, \dots, \alpha_i + \rho, \alpha_{i+1}, \dots, \alpha_n$, 且

$$B^T B v_i = (A^T A + 2\rho \alpha_i v_i v_i^T + \rho^2 v_i v_i^T) v_i =$$

$$A^T A v_i + 2\rho \alpha_i v_i + \rho^2 v_i =$$

$$\alpha_i^2 v_i + 2\rho \alpha_i v_i + \rho^2 v_i =$$

$$(\alpha_i^2 + 2\rho \alpha_i + \rho^2) v_i,$$

上式说明矩阵 $B^T B$ 的对应于特征值 $\alpha_i^2 + 2\rho \alpha_i + \rho^2$ 的特征向量为 v_i , 因此矩阵 B 的对应于奇异值 $\alpha_i + \rho$ 的右奇异向量为 v_i .

因为左奇异向量 u_i 为矩阵 BB^T 的特征向量, 首先计算

$$BB^T = (A + \rho u_i v_i^T)(A + \rho u_i v_i^T)^T =$$

$$AA^T + \rho A v_i u_i^T + \rho u_i v_i^T A^T + \rho^2 v_i^T v_i u_i u_i^T =$$

$$AA^T + \rho U \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix} V^T v_i u_i^T +$$

$$\rho u_i v_i^T V \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix}^T U^T + \rho^2 u_i u_i^T =$$

$$AA^T + \rho U \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix} e_i u_i^T +$$

$$\rho u_i e_i^T \begin{bmatrix} \sum & 0 \\ 0 & 0 \end{bmatrix}^T U^T + \rho^2 u_i u_i^T =$$

$$AA^T + 2\rho \alpha_i u_i u_i^T + \rho^2 u_i u_i^T,$$

于是

$$BB^T u_i = (AA^T + 2\rho \alpha_i u_i u_i^T + \rho^2 u_i u_i^T) u_i =$$

$$AA^T u_i + 2\rho \alpha_i u_i + \rho^2 u_i =$$

$$\alpha_i^2 u_i + 2\rho \alpha_i u_i + \rho^2 u_i =$$

$$(\alpha_i^2 + 2\rho \alpha_i + \rho^2) u_i,$$

即矩阵 BB^T 的对应于特征值 $\alpha_i^2 + 2\rho \alpha_i + \rho^2$ 的特征向量为 u_i , 因此矩阵 B 的对应于奇异值 $\alpha_i + \rho$ 的左奇异向量为 u_i , 所以矩阵 A 的左奇异向量仍为矩阵 $A + \rho u_i v_i^T$ 的左奇异向量.

定理 2 设 $x \in \mathbf{R}^m, y \in \mathbf{R}^n$, 令 $A = xy^T, H_l = (h_{l1}, h_{l2}, \dots, h_{lm}) \in \mathbf{R}^{m \times m}, H_r = (h_{r1}, h_{r2}, \dots, h_{rn}) \in \mathbf{R}^{n \times n}$ 为 Householder 矩阵, 使得

$$H_l x = \|x\|_2 e_1, H_r y = \|y\|_2 e_1,$$

则

- (i) H_l 的所有列构成 A 的左奇异向量基,
- (ii) H_r 的所有列构成 A 的右奇异向量基,
- (iii) $A = \|x\|_2 \|y\|_2 h_{l1} h_{r1}^T$.

证明 显然矩阵 $A = xy^T$ 的奇异值为 $\|x\|_2 \|y\|_2, 0, \dots, 0$, 对应于奇异值为 $\|x\|_2 \|y\|_2$ 的左、右奇异向量分别为 $\frac{x}{\|x\|_2}, \frac{y}{\|y\|_2}$,

$$H_l A H_r^T H_l x y^T H_r^T = \begin{bmatrix} \|x\|_2 \|y\|_2 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}_{m \times n},$$

所以 H_l 的所有列构成 A 的左奇异向量基, H_r 的所有列构成 A 的右奇异向量基.

$$\text{由 } H_l x = \|x\|_2 e_1 \text{ 得}$$

$$x = \|x\|_2 H_l e_1 = \|x\|_2 h_{l1},$$

$$\text{由 } H_r y = \|y\|_2 e_1 \text{ 得}$$

$$y = \|y\|_2 H_r e_1 = \|y\|_2 h_{r1},$$

$$\text{因此 } A = xy^T = \|x\|_2 \|y\|_2 h_{l1} h_{r1}^T.$$

令 $A = xy^T, H_l = (h_{l1}, h_{l2}, \dots, h_{lm}) \in \mathbf{R}^{m \times m}, H_r = (h_{r1}, h_{r2}, \dots, h_{rn}) \in \mathbf{R}^{n \times n}$ 为 Householder 矩阵, 使得 $H_l x = \|x\|_2 e_1, H_r y = \|y\|_2 e_1$, 由定理 1 及定理 2 知可对矩阵 A 进行一步一步的秩 1 修正, 使得 $\alpha_1, \alpha_2, \dots, \alpha_n$ 为所得矩阵的奇异值, 且 x, y 分别为其左右奇异向量, 具体步骤如下:

$$A = xy^T, \text{ 奇异值为 } \|x\|_2 \|y\|_2, 0, \dots, 0 (n-1 \text{ 个}),$$

↓ 进行秩 1 修正

$$A_1 = A + (\alpha_1 - \|x\|_2 \|y\|_2) h_{l1} h_{r1}^T, \text{ 由定理 1 知 } A_1 \text{ 的}$$

奇异值为 $\sigma_1, 0, \dots, 0(n-1$ 个),

↓ 进行秩 1 修正

$$A_2 = A + (\sigma_1 - \|x\|_2 \|y\|_2) h_1 h_1^T + \sigma_2 h_2 h_2^T, \text{ 奇异值为}$$

$\sigma_1, \sigma_2, 0, \dots, 0(n-2$ 个),

⋮

$$A_n = A + (\sigma_n - \|x\|_2 \|y\|_2) h_n h_n^T + \sum_{i=2}^n \sigma_i h_i h_i^T, \text{ 奇异}$$

值为 $\sigma_1, \sigma_2, \dots, \sigma_n$.

再由定理 2 可得

$$A_n = A + (\sigma_1 - \|x\|_2 \|y\|_2) h_1 h_1^T +$$

$$\sum_{i=2}^n \sigma_i h_i h_i^T = \sum_{i=1}^n \sigma_i h_i h_i^T.$$

由此得

算法:

- 输入 $x, y, \sigma_1, \sigma_2, \dots, \sigma_n$;
- 令 $A = 0$;
- 计算 Householder 阵 H_l, H_r , 使得 $Hx = \|x\|_2 e_1$;
 $H_r y = \|y\|_2 e_1$;
- $i = 1, 2, \dots, n$,
- $A = A + \sigma_i h_i h_i^T$;
- 输出 A .

上述算法只需要计算 2 个向量加法, $n+6$ 个向量乘法, $n+1$ 个矩阵加法, 因此运算量少并且易于进行并行计算.

定理 3 矩阵 $h_1 h_1^T, h_2 h_2^T, \dots, h_n h_n^T$ 线性无关.

证明 假设存在 n 个实数 k_1, k_2, \dots, k_n , 使得

$$k_1 h_1 h_1^T + k_2 h_2 h_2^T + \dots + k_n h_n h_n^T = \mathbf{0},$$

则等式两端同时右乘 $h_i (i = 1, 2, \dots, n)$ 得

$$(k_1 h_1 h_1^T + k_2 h_2 h_2^T + \dots + k_n h_n h_n^T) h_i = \mathbf{0},$$

所以

$$k_i h_i = \mathbf{0},$$

即 $k_i = 0 (i = 1, 2, \dots, n)$, 说明 $h_1 h_1^T, h_2 h_2^T, \dots, h_n h_n^T$ 线性无关.

定理 4 $\sigma_i \neq \sigma_j (i \neq j)$ 时, 问题至少有 $n!$ 个解.

证明 在算法中改变 σ 的次序得

$$A_{\pi} = \sum_{i=1}^n \sigma_{\pi(i)} h_i h_i^T,$$

其中 $\{\sigma_{\pi(1)}, \sigma_{\pi(2)}, \dots, \sigma_{\pi(n)}\}$ 为 $\{\sigma_1, \sigma_2, \dots, \sigma_n\}$ 的一个置换, 因此共有 $n!$ 种不同的置换. 由定理 3 知 $h_1 h_1^T, h_2 h_2^T, \dots, h_n h_n^T$ 线性无关, 因此 Householder 阵 H_l, H_r 确定后, 当 $\sigma_i \neq \sigma_j (i \neq j)$ 时, 问题 I 至少有 $n!$ 个解.

注:

1) 取 $w = x \pm \|x\|_2 e_1$, 令 Householder 阵 $H_l = I - 2ww^T / (w^T w)$, 则 $H_l x = \|x\|_2 e_1$. 因此 Householder 阵有

两种选择.

2) 当算法中 Householder 阵 H_l, H_r 确定后, 不论怎样改变 σ 的次序, 得到的矩阵 A 在相差一置换阵的条件下可以认为是唯一的, 即存在置换阵 P , 使得 $A = PA_{\pi} P^T$.

在式(1)中令 $A_0 = 0, A_i = h_i h_i^T$, 即为本文中讨论的逆奇异值问题. 若取 $m = n, x = y$, 按照本文的算法可构造 n 阶对称阵 A 具有给定的 n 个特征值及特征向量 x , 即文献[14]的情况.

2 数值例子

下面给出数值例子来验证上述结论.

例 1 令 $m = 8, n = 6$, 选取一个随机矩阵 $A_0 \in \mathbf{R}^{m \times n}$:

$A_0 =$

0.7709	0.1131	0.7540	0.8506	0.4881	0.6737
0.3139	0.8121	0.6632	0.3402	0.9926	0.9573
0.6382	0.1564	0.2722	0.9138	0.5314	0.1112
0.9866	0.1564	0.2722	0.9138	0.5314	0.1112
0.5029	0.1221	0.4194	0.2286	0.1813	0.5651
0.9477	0.7627	0.2130	0.8620	0.5019	0.9692
0.8280	0.7218	0.0356	0.6566	0.4222	0.0237
0.9176	0.6516	0.0812	0.8912	0.6604	0.8702

可求得 A_0 的奇异值为 4.078 1, 1.164 3, 0.908 6, 0.855 1, 0.472 0, 0.187 4, 其中最大奇异值 4.078 1 对应的左右奇异向量分别记为 x, y , 则

$$x = [-0.3664, -0.3843, -0.3314, -0.3204, -0.2014, -0.4441, -0.2894, -0.4290]^T,$$

$$y = [-0.5115, -0.3816, -0.2784, -0.4637, -0.3725, -0.4018]^T.$$

取给定的奇异值 $\sigma_1, \sigma_2, \dots, \sigma_n$ 为 A_0 的奇异值, 对上述给定的 x 和 y , 由算法得所求的矩阵为

$A_0 =$

1.2148	0.2366	0.1979	0.5571	0.5726	0.6368
0.4841	1.5564	0.2931	0.5522	0.4788	0.5218
0.5476	0.3710	1.1853	0.5161	0.4450	0.4845
0.3775	0.3203	0.2396	1.3074	0.3927	0.4281
0.3106	0.2198	0.1641	0.3068	0.7369	0.2886
0.9976	0.5638	0.4195	0.7726	0.6613	0.9069
0.6991	0.3797	0.2824	0.5185	0.4430	0.4819
1.0364	0.5629	0.4186	0.7686	0.6567	0.7143

取 $10^{-10} \leq \varepsilon \leq 10^{10}$, 现在给奇异值一个扰动得到 $\sigma^{\varepsilon} = \sigma_i + \varepsilon$, 运用算法, 我们得到对应于 σ^{ε} 的 A^{ε} , 从图 1 可以看出随着 ε 趋近于零, A^{ε} 趋近于 A .

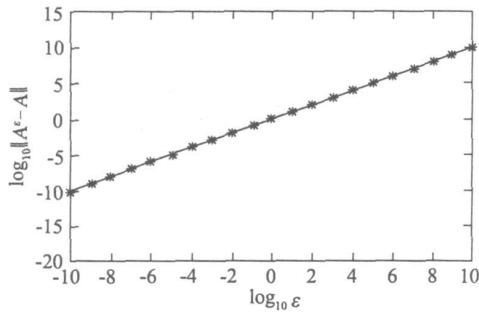


图1 算法误差曲线

Fig. 1 Curve of algorithmic error

参考文献:

- [1] Gramm A G. Inverse problems: mathematical and analytical techniques with applications to engineering[M]. German: Springer, 2005.
- [2] Lebedev L P, Vorovich I I, Gladwell G M. Functional analysis: applications in mechanics and inverse problems[M]. Netherland: Kluwer Academic Publishers, 2002.
- [3] Uhlmann G, Levy S. Inside out: inverse problems and applications[M]. Cambridge: Cambridge University Press, 2003.
- [4] Chu M T. Inverse eigenvalue problems[J]. SIAM Rev, 1998, 40: 1-39.
- [5] Chu M T, Golub G H. Structured inverse eigenvalue problems[J]. Acta Numer, 2002, 11: 1-71.
- [6] Chu M T, Golub G H. Inverse eigenvalue problems: theory, algorithms, and applications[M]. England: Oxford University Press, 2005.
- [7] Xu S F. An introduction to inverse eigenvalue problems[M]. Beijing: Peking University Press and Vieweg Publishing, 1998.
- [8] 周树荃, 戴华. 代数特征值反问题[M]. 郑州: 河南科学技术出版社, 1991.
- [9] Chu M T. Numerical methods for inverse singular value problems[J]. SIAM J Numer Anal, 1992, 29: 885-903.
- [10] Friedland S, Nocedal J, Overton M L. The formulation and analysis of numerical methods for inverse eigenvalue problems[J]. SIAM J Numer Anal, 1987, 24: 634-667.
- [11] Bai Z J, Xu S F. An inexact Newton-type method for inverse singular value problems[J]. Lin Alg Appl, 2008, 429: 527-547.
- [12] Politi T. A discrete approach for the inverse singular value problem in some quadratic group[J]. Lecture Notes in Comput Sci, 2003, 2658: 121-130.
- [13] Chu D, Chu M T. Low rank update of singular values[J]. Math Comp, 2006, 75: 1351-1366.
- [14] Zhu Shengxin, Gu Tongxiang, Liu Xingping. Solving inverse eigenvalue problems via Householder and rank-one matrices[J]. Lin Alg Appl, 2009, 430(1): 318-334.
- [15] Golub G H, Van Loan C F. Matrix computations[M]. American: The Johns Hopkins University Press, 1996.
- [16] 袁亚湘. 矩阵计算[M]. 北京: 科学出版社, 2001.
- [17] Brauer A. Limits for the characteristic roots of a matrix IV. Applications to stochastic matrices[J]. Duke Math J, 1952, 19: 75-91.

A Special Kind of Inverse Singular Value Problem of Matrices

WU Chun-hong^{*}, LU Lin-zhang^{**}

(School of Mathematical Science, Xiamen University, Xiamen 361005, China)

Abstract: In this paper, we consider a kind of matrix inverse singular value problem. Given nonnegative numbers $\sigma_1, \sigma_2, \dots, \sigma_n$, two nonzero real vectors $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$, $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$, find $m \times n$ real matrix \mathbf{A} , such that $\sigma_1, \sigma_2, \dots, \sigma_n$ are the singular values of \mathbf{A} , and \mathbf{x}, \mathbf{y} are the left and right singular vectors, respectively. Based on Householder transformation and rank-one updating matrices, we propose an algorithm which is economical and easily to parallel to solve the inverse singular value problem. We also give the corresponding numerical example.

Key words: inverse problem; singular value; Householder transformation; rank-one updating