



On singular values of partially prescribed matrices[☆]

Pedro M.Q. Aguiar, Marko Stošić*, João Xavier

Instituto de Sistemas e Robótica, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal

Received 31 August 2007; accepted 9 June 2008

Available online 3 August 2008

Submitted by J.F. Queiró

Abstract

In this paper we study singular values of a matrix whose one entry varies while all other entries are prescribed. In particular, we find the possible p th singular value of such a matrix, and we define explicitly the unknown entry such that the completed matrix has the minimal possible p th singular value. This in turn determines possible p th singular value of a matrix under rank one perturbation. Moreover, we determine the possible value of p th singular value of a partially prescribed matrix whose set of unknown entries has a form of a Young diagram. In particular, we give a fast algorithm for defining the completion that minimizes the p th singular value of such matrix.

© 2008 Elsevier Inc. All rights reserved.

AMS classification: 15A18

Keywords: Singular value; Matrix completion; Rank deficiency matrix

1. Introduction

In this paper we are interested in completions of a partially prescribed matrix, such that the rank of the completed matrix is as small as possible, and in defining the completion when this minimum is obtained.

[☆] This work was partially supported by the Portuguese Foundation for Science and Technology (FCT), under ISR/IST plurianual funding (POSC program, FEDER), and grants MODI-PTDC/EEA-ACR/72201/2006 and SIPM-PTDC/EEA-ACR/73749/2006.

* Corresponding author.

E-mail addresses: aguiar@isr.ist.utl.pt (P.M.Q. Aguiar), mstosic@isr.ist.utl.pt (M. Stošić), jxavier@isr.ist.utl.pt (J. Xavier).

Problems of this kind have been considered previously, e.g. in the papers by Cohen et al. [4], Gohberg et al. [6], Rodman and Woerdeman [11].

Moreover, direct motivation for studying these problems comes from computer vision. Several computer vision tasks have been approached by using methods that require finding linear or affine low dimensional subspaces from noisy observations. Usually, those subspaces are found by computing rank deficient matrices from noisy observations of their entries. Examples of these tasks include applications in photogrammetry [2], image registration [16], object recognition [15], and the construction of 3D models from 2D images [14,1].

An alternative way of saying that a matrix has the rank at most r is to say that its $(r + 1)$ th singular value is equal to 0. Thus, the problem becomes to complete a partially prescribed matrix such that its $(r + 1)$ th singular value is equal to 0. This formulation has the advantage in applications, where matrices are usually corrupted with noise. In that case, it should be required that the completed matrix is close to a rank r matrix (in terms of, e.g., Frobenius or spectral distance), i.e. that the $(r + 1)$ th singular value is as small as possible.

Hence, the most natural generalization of this problem is

Problem 1. Determine the possible values of p th singular value of a partially prescribed matrix.

Related problems of determining the possible singular values of the sum of two matrices have been studied by Queiró and de Sá [9] (see also the review paper by Fulton [5]).

In this paper we solve Problem 1 in the case when the set of unknown entries has the form of a Young diagram, which includes the cases when the set of unknown entries has the form of a submatrix, or a triangle, which is of high interest in the applications.

We give complete solution for the generic matrix, and for an arbitrary matrix we show that we can complete it such that the resulting matrix has the p th singular value arbitrarily close to a theoretical minimum. Moreover, we give a fast and efficient algorithm for the definition of the unknown entries such that the completed matrix has the prescribed singular value.

The solution of Problem 1 is split in two problems, solved in Sections 3 and 4:

Consider the matrix whose only one entry varies while all other are prescribed. Denote the value of the unprescribed entry by x . In Section 3, Theorems 2 and 5, we solve the following problem:

Problem 2. Find the possible values of σ_p , when x varies, and find x when the minimum is obtained.

Problem 2 can be equivalently stated as the problem of determining the possible p th singular value of a given matrix under specified rank one perturbation. In other words, if $A \in \mathbb{R}^{m \times n}$ is the prescribed matrix, and $f \in \mathbb{R}^{m \times 1}$ and $g \in \mathbb{R}^{n \times 1}$ are nonzero vectors, in Section 3 we determine

$$\inf_{t \in \mathbb{R}} \sigma_p(A + tfg^T) \quad \text{and} \quad \sup_{t \in \mathbb{R}} \sigma_p(A + tfg^T). \quad (1)$$

The related problem of the prescription of the possible singular values of a matrix when an arbitrary rank one perturbation is performed, was solved in the case of square matrices by Thompson in [12]. In that case, the vectors f and g are also allowed to vary, and that gives much more freedom. In the problem we are interested in, the vectors f and g are fixed (or equivalently, all entries of an arbitrary rectangular matrix A except one are fixed) which makes the problem more complicated. Moreover, we give explicit definition of the unique unknown entry when the extremal values from (1) are obtained.

In Section 4, we study singular values of a real matrix M of dimension $n \times m$, with the property that if the entry (i, j) is known then all the entries (k, l) , with $k \geq i$ and $l \geq j$ in M , are known, i.e. when the set of unknown entries has the form of a Young diagram. Particular cases are when the unknown entries form a submatrix or a triangle. In Theorem 7, we solve the following problem:

Problem 3. Find the possible values of p th singular value of M , and define the completion when the minimum is obtained.

As we said previously, our major motivation for these problems comes from computer vision. In the factorization method for the recovery of 3D rigid structure from video, a set of points is tracked across time and their trajectories are collected in an observation matrix, where each column contains the sequence of coordinates of each feature point. Due to the rigidity of the scene, the observation matrix is rank deficient in a noiseless situation, see, e.g. [14,1]. When the observation matrix is completely prescribed, the solution to the problem of finding the nearest matrix with prescribed rank is easily obtained from its Singular Value Decomposition. However, in practice, the observation matrix may be only partially prescribed, because some of its entries may be unknown (unobserved). Various similar problems have been previously studied by using numerical optimization methods. For example, by developing sub-optimal solutions to combine the constrains that arise from the observed submatrices of the original matrix [14,8], or by minimizing the nonlinear cost function that accounts for the residual of the approximation [7,3]. As usual in nonconvex optimization, these approaches lead to iterative algorithms whose convergence to the global minimizers is highly dependent on the initialization.

In practice, when a feature point is occluded (or missed during tracking for any other reason) at a given frame, the corresponding trajectory ends at that frame and the remaining entries of the corresponding column of the observation matrix must be treated as missing data. In these situations, by re-arranging the observation matrix, through column re-ordering, the pattern of missing entries becomes a Young diagram.

In this paper, we show how to complete partially prescribed matrices in an optimal way, when the set of unknown entries has the form of a Young diagram, by using purely linear algebra methods.

2. Preliminary results and notation

Let $A \in \mathbb{R}^{m \times n}$. We shall assume that $n \geq m$. If $n < m$, the singular values of A we define to be the singular values of the matrix A^T .

By *singular values* of the matrix A we mean the nonnegative real numbers $\sigma_1 \geq \dots \geq \sigma_m \geq 0$ such that $\sigma_1^2 \geq \dots \geq \sigma_m^2$ are the eigenvalues of $AA^T \in \mathbb{R}^{m \times m}$. Also, we assume that $\sigma_i = +\infty$ for $i \leq 0$, and $\sigma_i = 0$ for $i > m$.

Then, there exist orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$, such that

$$UAV = \begin{bmatrix} \sigma_1 & \cdots & 0 & 0 \\ & \ddots & & 0 \\ 0 & \cdots & \sigma_m & 0 \end{bmatrix}. \tag{2}$$

The form (2) of the matrix A is called *the singular value decomposition (SVD)*. Moreover, from (2), we have that multiplying the matrix A from left and from right by orthogonal matrices does not change its singular values. In particular, permuting rows or permuting columns of a matrix does not change its singular values.

If $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_m = 0$, then $\text{rank } A = r$. Moreover if $\sigma_{r+1} = 0$, then $\text{rank } A \leq r$.

Let M_r be the set of all m by n real matrices, with the rank at most r . Then by SVD (2) we have that the (Frobenius) distance from matrix A to the set M_r is given by

$$\text{dist}(A, M_r) = \sqrt{\sum_{i=r+1}^m \sigma_i^2} \leq \sqrt{m-r} \sigma_{r+1}.$$

Moreover, a matrix from M_r , closest to A is

$$A_r = U \begin{bmatrix} \sigma_1 & \dots & 0 & 0 \\ & \ddots & & \\ 0 & \dots & \sigma_r & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix} V.$$

Thus, we have that σ_{r+1} is small (less than given positive tolerance ϵ), if and only if, A is close to the rank r matrix, and in this case, we say that the *numerical rank* of A is at most r .

2.1. Interlacing inequalities

Interlacing inequalities present the relation between the singular values of a matrix and its submatrix. For details, see [13] or [10]

Theorem 1. Let $A \in \mathbb{R}^{n \times m}$, $u \in \mathbb{R}^{n \times 1}$ and

$$B = \begin{bmatrix} A & u \end{bmatrix} \in \mathbb{R}^{n \times (m+1)}.$$

Let $s_1 \geq s_2 \geq \dots$ be singular values of A , and $s'_1 \geq s'_2 \geq \dots$ be singular values of B . Then

$$s'_1 \geq s_1 \geq s'_2 \geq s_2 \geq s'_3 \geq s_3 \geq \dots$$

3. One unknown entry

In this section we determine the possible p th singular value of a partially prescribed matrix whose only one entry is unknown. By permuting rows and columns, we can assume that the unknown entry is at the position (1,1).

Let $u \in \mathbb{R}^{k \times 1}$, $v \in \mathbb{R}^{n \times 1}$ and $A \in \mathbb{R}^{k \times n}$ be prescribed matrices. Consider the matrix

$$M = \left[\begin{array}{c|c} x & v^T \\ \hline u & A \end{array} \right] \in \mathbb{R}^{(k+1) \times (n+1)}, \tag{3}$$

where $x \in \mathbb{R}$ varies. Without loss of generality, we can assume that $k \leq n$. Denote the singular values of (3) by $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{k+1}$.

Moreover, denote by $s_1 \geq s_2 \geq \dots \geq s_k$ and $s'_1 \geq s'_2 \geq \dots \geq s'_{k+1}$ the singular values of the matrices

$$B = \begin{bmatrix} u & A \end{bmatrix} \in \mathbb{R}^{k \times (n+1)}$$

and

$$C = \begin{bmatrix} v^T \\ A \end{bmatrix} \in \mathbb{R}^{(k+1) \times n},$$

respectively, and let $\bar{s}_1 \geq \bar{s}_2 \geq \dots \geq \bar{s}_k$ be singular values of the matrix A .

Then, by interlacing inequalities (Theorem 1), we have

$$\max\{s_p, s'_p\} \leq \sigma_p \leq \min\{s_{p-1}, s'_{p-1}\}. \tag{4}$$

Thus, $\max\{s_p, s'_p\}$ is the lower bound for σ_p . In the following theorem, we prove that for generic matrix M (by the generic set of matrices we mean any subset S of all matrices, such that the closure of S in Frobenius metric is equal to the whole $\mathbb{R}^{(k+1) \times (n+1)}$), any value from the interval (4) can be obtained as the p th singular value of matrix M , when x varies.

Theorem 2. *Suppose that the numbers $s_p, s'_p, s_{p-1}, s'_{p-1}, \bar{s}_p$ and \bar{s}_{p-1} are all pairwise distinct. Let $s \in \mathbb{R}$ be a number such that*

$$\max\{s_p, s'_p\} \leq s \leq \min\{s_{p-1}, s'_{p-1}\}.$$

Then there exists x , such that the matrix (3) has s as the p th singular value.

Moreover, there exists a unique x such that the matrix (3) has $\max\{s_p, s'_p\}$ as the p th singular value.

Proof. Since the singular values change continuously with the change of the entries of the matrix, it is enough to show that there exists x such that the matrix M has the p th singular value equal to $\max\{s_p, s'_p\}$, and that there exists some other value of x for which the matrix M has the p th singular value equal to $\min\{s_{p-1}, s'_{p-1}\}$. We shall show how to define x in the first case – the second one can be obtained analogously.

Without loss of generality, we assume that $s_p > s'_p$. In order to prove that the p th singular value of matrix M is equal to s_p , note that it is enough to show that s_p is a singular value of M . Indeed, from:

$$\sigma_p \geq s_p > s'_p \geq \sigma_{p+1},$$

we conclude that then $\sigma_p = s_p$. Hence, we are left with defining x such that the matrix M has s_p as a singular value.

From the definition of singular values, we have that s_p is a singular value of (3), if and only if s_p^2 is an eigenvalue of MM^T , i.e. if

$$\det(MM^T - s_p^2 I) = 0. \tag{5}$$

From (5) and (3), we have

$$p(x) = \det \begin{bmatrix} x^2 + v^T v - s_p^2 & x u^T + v^T A^T \\ x u + A v & u u^T + A A^T - s_p^2 I \end{bmatrix} = 0. \tag{6}$$

In the matrix in (6), we have x^2 only at the position (1,1), and linear factors of x in the first row and first column. Thus, $p(x)$ is a quadratic function of x , i.e.,

$$p(x) = ax^2 + bx + c \tag{7}$$

for some $a, b, c \in \mathbb{R}$. We shall show below that the discriminant $D = b^2 - 4ac$ is equal to 0, and that $a \neq 0$, and so that there exists unique x , given by

$$x = -\frac{b}{2a}$$

for which the matrix (3) has s_p as singular value, as wanted.

Thus, we are left with showing that $b^2 = 4ac$ and $a \neq 0$.

Recall that by $s_1 \geq s_2 \geq \dots$ we have denoted the singular values of the matrix $B = [u \ A]$. Since

$$[u \ A] \begin{bmatrix} u^T \\ A^T \end{bmatrix} = uu^T + AA^T,$$

we have

$$\det(uu^T + AA^T - s_p^2 I) = 0.$$

We define $N = uu^T + AA^T - s_p^2 I$. Thus we have $\det N = 0$ and $N = N^T$.

By using this notation, and the fact that the determinant of a matrix is the linear function of its columns and rows, Eq. (6) becomes

$$\begin{aligned} p(x) &= \det(MM^T - s_p^2 I) = \det \begin{bmatrix} x^2 + v^T v - s_p^2 & xu^T + v^T A^T \\ xu + Av & N \end{bmatrix} \\ &= \det \begin{bmatrix} x^2 & xu^T + v^T A^T \\ xu & N \end{bmatrix} + \det \begin{bmatrix} v^T v - s_p^2 & xu^T + v^T A^T \\ Av & N \end{bmatrix} \\ &= \det \begin{bmatrix} x^2 & xu^T \\ xu & N \end{bmatrix} + \det \begin{bmatrix} 0 & v^T A^T \\ xu & N \end{bmatrix} \\ &\quad + \det \begin{bmatrix} 0 & xu^T \\ Av & N \end{bmatrix} + \det \begin{bmatrix} v^T v - s_p^2 & v^T A^T \\ Av & N \end{bmatrix} \\ &= x^2 \det \begin{bmatrix} 0 & u^T \\ u & N \end{bmatrix} + 2x \det \begin{bmatrix} 0 & v^T A^T \\ u & N \end{bmatrix} + \det \begin{bmatrix} 0 & v^T A^T \\ Av & N \end{bmatrix} \\ &= ax^2 + bx + c. \end{aligned}$$

Here, in the third equality we have used the fact that $\det N = 0$.

Since N is the symmetric matrix and $\det N = 0$, there exists orthogonal matrix $P \in \mathbb{R}^{k \times k}$ such that

$$PNP^T = \begin{bmatrix} b_1 & 0 & \dots & 0 & 0 \\ 0 & b_2 & \dots & 0 & 0 \\ \dots & \dots & \ddots & 0 & 0 \\ 0 & 0 & \dots & b_{k-1} & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Let $\beta = b_1 b_2 \dots b_{k-1}$ and

$$Pu = \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix}, \quad PA v = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}.$$

Then, we have

$$a = -\beta u_k^2, \quad b = -2\beta u_k v_k, \quad c = -\beta v_k^2, \tag{8}$$

which gives $b^2 = 4ac$ as wanted.

Finally, we are left with proving that $a \neq 0$.

Suppose that $a = 0$. Then, from (8), we would have that the matrix $[u \ N] \in \mathbb{R}^{k \times (n+1)}$ has the rank strictly less than k . Moreover, since $N = uu^T + AA^T - s_p^2 I$, the matrix $[u \ N]$ is

equivalent to $[u \quad AA^T - s_p^2 I]$, and so the matrix $AA^T - s_p^2 I$ cannot be of full rank. In other words,

$$\det(AA^T - s_p^2 I) = 0$$

and so s_p is a singular value of A .

However, from the interlacing inequalities, we have

$$\bar{s}_{p-1} \geq s_p \geq \bar{s}_p$$

and so either \bar{s}_{p-1} or \bar{s}_p must be equal to s_p , which is a contradiction. Thus, we have $a \neq 0$, which concludes our proof. \square

Remark 3. Although we are using SVD forms in the proof, to actually obtain the value x for which the theoretical minimum for σ_p is attained, we do not need them.

In fact, by replacing x with 0, 1 and -1 in (6), we can obtain the values $p(0)$, $p(1)$ and $p(-1)$ as the determinants of $(k + 1) \times (k + 1)$ matrix:

$$p(i) = \det \begin{bmatrix} i^2 + v^T v - s_p^2 & iu^T + v^T A^T \\ iu + Av & uu^T + AA^T - s_p^2 I \end{bmatrix}, \quad i \in \{-1, 0, 1\}.$$

Thus, the coefficients a , b and c from (7) are

$$a = \frac{p(1) + p(-1)}{2} - p(0), \tag{9}$$

$$b = \frac{p(1) - p(-1)}{2}, \tag{10}$$

$$c = p(0). \tag{11}$$

Since we have shown that $b^2 - 4ac = 0$ and $a \neq 0$, x is given by

$$x = -\frac{b}{2a}. \tag{12}$$

Remark 4. Note that in the course of the proof, we have obtained that in the generic case when all s_i , s'_i and \bar{s}_i are pairwise distinct, each of the numbers s_i and s'_i can be obtained as the singular value of the matrix M .

In the general case, it may be impossible to define x such that the minimum or maximum value for σ_p is attained. However, since every matrix can be approximated arbitrarily well (in the Frobenius metric) by a matrix that satisfies the conditions from Theorem 2, we obtain the following

Theorem 5. *Let s be arbitrary number such that*

$$\max\{s_p, s'_p\} < s < \min\{s_{p-1}, s'_{p-1}\}.$$

Then there exists $x \in \mathbb{R}$ such that the matrix M has s as the p th singular value.

In particular, the theoretical minimum $\max\{s_p, s'_p\}$ can be approximated with arbitrary precision.

Proof. Let $\epsilon > 0$ be arbitrary positive real number. If matrices A , u and v from (3) do not satisfy the conditions from Theorem 2, then there exist matrices A' , u' and v' , such that their distances

$$\max_{j=1}^{k+1} s_p(M(j, i_j + 1)) \leq \sigma_p \leq \min_{j=1}^{k+1} s_{p+1-j-i_j}(M(j, i_j + 1)). \quad (15)$$

Conversely, if s is such that

$$\max_{j=1}^{k+1} s_p(M(j, i_j + 1)) < s < \min_{j=1}^{k+1} s_{p+1-j-i_j}(M(j, i_j + 1)), \quad (16)$$

then there exists a completion of M , such that $\sigma_p = s$.

Proof. By interlacing inequalities, we have that for every $j = 1, \dots, k + 1$:

$$\sigma_p \geq s_p(M(j, i_j + 1))$$

and also that

$$\sigma_p \leq s_{p+1-j-i_j}(M(j, i_j + 1)),$$

which gives (15).

For the converse, we shall prove that both ends of the interval in (16) can be numerically reached (a value w can be numerically reached if for every $\epsilon > 0$, there exists a completion of M such that difference between $s_p(M)$ and w is less than ϵ). In particular, we shall prove this for the lower bound from (16), while the upper bound can be obtained completely analogously.

To this end, by repeated use of the Corollary 5, we shall obtain the matrix M such that its p th singular value attains numerically the lower bound. Indeed, choose any unknown entry (i, j) such that its value is unknown, while the entries $(i + 1, j)$ and $(i, j + 1)$ are known (these are the “corners” of the Young diagram, and in the example (14), these entries are denoted by y). Then define the value of that entry to be x , where x is such that $s_p(M(i, j))$ is numerically equal to $\max\{s_p(M(i + 1, j)), s_p(M(i, j + 1))\}$, as in the previous section.

Now, proceed with the process until all entries of the matrix are filled, and such that it has the wanted p th singular value. \square

References

- [1] P. Aguiar, J. Moura, Rank 1 weighted factorization for 3D structure recovery: algorithms and performance analysis, *IEEE Trans. Pattern Anal. Machine Intell.* 25 (9) (2003).
- [2] P. Belhumeur, D. Kriegman, What is the set of images of an object under all possible lighting conditions?, *Proc. IEEE Conf. Computer Vision Pattern Recognition* (1996) 270–277.
- [3] A. Buchanan, A. Fitzgibbon, Damped Newton algorithms for matrix factorization with missing data, *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, St Barbara, USA (2005).
- [4] N. Cohen, C.R. Johnson, L. Rodman, H. Woerdeman, Ranks of completions of partial matrices, *The Gohberg Anniversary Collection*, vol. I (Calgary, AB, 1988), pp. 165–185 (*Oper. Theory Adv. Appl.*, vol. 40, Birkhäuser, Basel, 1989).
- [5] W. Fulton, Eigenvalues, invariant factors, highest weights, and Schubert calculus, *Bull. Amer. Math. Soc.* 37 (3) (2000) 209–249.
- [6] I. Gohberg, L. Rodman, T. Shalom, H. Woerdeman, Bounds for eigenvalues and singular values of matrix completions, *Linear and Multilinear Algebra* 33 (3-4) (1993) 233–249.
- [7] R. Guerreiro, P. Aguiar, Estimation of rank deficient matrices from partial observations: two-step iterative algorithms, energy min. meth., *Computer Vision and Pattern Recognition*, Lecture Notes in Computer Science 2683, Springer-Verlag, 2003.
- [8] D. Jacobs, Linear fitting with missing data: applications to structure-from-motion and to characterizing intensity images, *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, St Barbara, USA, 1997.
- [9] J.F. Queiró, E.M. de Sá, Singular values and invariant factors of matrix sums and products, *Linear Algebra Appl.* 225 (1995) 43–56.

- [10] J.F. Queiró, On the interlacing property for singular values and eigenvalues, *Linear Algebra Appl.* 97 (1987) 23–28.
- [11] L. Rodman, H. Woerdeman, Perturbation, singular values, and ranks of partial triangular matrices, *SIAM J. Matrix Anal. Appl.* 16 (1) (1995) 278–288.
- [12] R.C. Thompson, The behaviour of eigenvalues and singular values under perturbations of restricted rank, *Linear Algebra Appl.* 13 (1976) 69–78.
- [13] R.C. Thompson, Principal submatrices IX: interlacing inequalities for singular values of submatrices, *Linear Algebra Appl.* 5 (1994) 1–12.
- [14] C. Tomasi, T. Kanade, Shape and motion from image streams under orthography: a factorization method, *Int. J. Comput. Vision* 9 (2) (1992) 137–154.
- [15] S. Ullman, R. Basri, Recognition by linear combinations of models, *IEEE Trans. Pattern Anal. Machine Intell.* 13 (10) (1991) 992–1007.
- [16] Lihi Zelnik-Manor, Michal Iranim, Multi-view subspace constraints on homographies, *Proc. of IEEE Int. Conf. on Computer Vision*, September, vol. 2, Kerkyra, Greece, 1999, pp. 710–715.