Numerical Analysis Project
Manuscript NA-86-36

October 1986

The truncated SVD as a method for regularization

by

Pet Christian Hansen

Numerical Analysis Project
Computer Science Department,
Stanford University
Stanford, California 94305

The ‘truncated SVD as a method for regularization

Per Christian Hansen*

Abstract

The truncated singular value decomposition (SVD) is considered as a method for regularization of ill-posed linear least squares problems. In particular, the truncated SVD solution is compared with the usual regularized solution. Necessary conditions are defined in which the two methods will yield similar results. This investigation suggests the truncated SVD as a favorable al-. ternative to standard-form regularization in case of ill-conditioned matrices with a well-determined rank.

Key words: truncated singular value decomposition, regularization in standard form, perturbation theory for truncated SVD, numerical rank.

1. Introduction

This paper deals with methods for solving the unconstrained linear least squares problem

$$\mathrm{niin}\,\|b - A\,x\|\ ,\quad A \in R^{m\times n}\ ,\quad m \geq n\ . \tag{1}$$

Here, and throughout the pppcr, $\|\,.\,\| = \|\cdot\|_2$. When the matrix A is ill-conditioned, the problem (1) is ill-posed in the sense that a small perturbation of b may lead to a large perturbation of the solution. The same is true for perturbations Of A. A well-known and highly regarded method for dealing with such ill-posed problems is the method of regularization by Tikhonov [18] and Phillips [I 7]. In particular. *regularization in standard form* corresponds to defining a regularized solution x_λ, as a function of the regularization parameter λ, by

$$x_A \equiv \mathrm{argmin}\,\{\|b - A\,x\|^2 + \lambda^2\,\|x\|^2\}\ . \tag{2}$$

It is easy to show that x_λ is the least squares solution to the problems

$$\min\left\|\begin{bmatrix} b \\ 0 \end{bmatrix} - \begin{bmatrix} A \\ \lambda I_n \end{bmatrix} x\right\| \tag{3}$$

where I_n denotes the identity matrix of order n, and x_λ is unique since the augmented matrix in (3) has full rank.

Another well-known method for dealing with ill-conditioned matrices in problem (1) is the *truncated singular value decomposition* (TSVD), cf. 1 Ianson [12] and Varah [21]. The use Of the TSVD has certain similarities with the user of regularization in standard from, and it is generally known that the two methods often produce very similar results [22]. The purpose of this paper is to investigate the connection between the two methods and define necessary conditions in which the two methods will yield similar results.

* On leave from Copenhagen University Observatory, Øster Voldgade 3, DK-1350 Copenhagen K, Denmark. This work was in part supported by National Science Foundation Grant Number DCR 8412314 and by a Fulbright Supplementary Grant.

The author is aware that it is often necessary to substitute $\| L\, x \|$ for $\| x \|$ in (2), and that this may have a considerable effect on the solution [25]. However, it is stressed that analysis of the standard-form problem (2) will shed light on some general relations between TSVD and regularization, and that a problem not in standard form can be transformed into a problem in standard form as shown by Eldén [6].

The organization of the paper is as follows. In Section 2 the TSVD and standard-form regularization are stated in terms of the SVD of the matrix A, and in Section 3 a new perturbation theory for the TSVD is given. On basis of this, it is natural to divide ill-conditioned matrices into two classes of matrices with well-determined and ill-determined numerical rank, respectively, as discussed in Section 4. Section 5 treats the case of matrices with well-determined rank, and it is verified that TSVD and regularization can, produce similar solutions. Finally, the case of matrices with ill-determined rank is treated in Section 6 where it is shown that the similarity of the two solutions now depends on the projection of the right-hand side 6 onto the left singular vectors of A.

The paper is similar in spirit to the papers of Varah [22] and Wedin [23,24], and extensions of the results of Wedin are given.

2. Truncated SVD and standard-form regularization

Throughout the paper, the singular value decomposition (SVD) of the matrix A in (l) will be extensively used. To summarize the SVD briefly, let A be decomposed into the three matrices U, Σ, and V:

$$A = U\, \Sigma\, V^T \tag{4}$$

where the left and right singular matrices $U \in R^{m\times m}$ and $V \in R^{n\times n}$ are orthogonal, and where the matrix $\Sigma \in R^{m\times n}$ has diagonal form:

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots., \sigma_n) \quad , \tag{5}$$

The diagonal elements $\{\sigma_i\}$ of Σ are the singular values of A, and they are ordered such that:

$$\sigma_1 \geq \sigma_2 \geq \ . \ . \ . \geq \sigma_r > \sigma_{r+1} = \ . \ . \ . = \sigma_n = 0 \tag{6}$$

where r = rank (A). In particular, $\| A \| = \sigma_1$. For a rigorous treatment of the SVD, see e.g. [LO]. It is stressed that the SVD is mainly used here as a powerful analysis tool. The TSVD and the regularized solutions as defined below can be computed with much less computational effort my means of other methods [4,6].

The basic idea of TSVD as wc!! as standard-form regularization is to impose the additional requirement on the solution that its norm be small, thus hopefully damping lllc contributions from the errors of the right-hand side. In the case of TSVD, this is achieved by neglection of the components of the solution corresponding to the smallest singular values, since these contributions lo the solution are most likely to be large. Thus, the TSVD of A is defined as the rank-k matrix

$$A_k \equiv U\, \Sigma_k\, V^T = \sum_{i=1}^{k} u_i\, \sigma_i\, v_i^T \quad , \quad \Sigma_k = \operatorname{diag}(\sigma_1, \ldots., \sigma_k, 0, \ldots.0) \in R^{m\times n} \tag{7}$$

where Σ_k equals Σ with the smallest n-k singular values replaced by zeroes, and $k \leq r$. u_i and v_i are the columns of the matrices U and V, respectively. When the number k is chosen properly, then the condition number σ_1/σ_k of the TSVD A_k will be small. The TSVD solution to (1), defined by:

$$x_k \equiv A_k^+\ 6 \ , \tag{8}$$

is therefore not very sensitive to errors in 6 and A. The matrix A_k^+ is the pseudoinverse of A_k:

$$A_k^+ = V \Sigma_k^+ U^T \quad , \quad \Sigma_k^+ = \mathrm{diag}(\sigma_1^{-1}, \dots, \sigma_k^{-1}, 0, \dots, 0) \in R^{n\times m} \quad , \tag{9}$$

and A_k^+ is actually a {2,3,4}-invcrsc, or outer invcrsc, of A, cf. [2]. Thc TSVD solution can usually bc computcd from a Q-R factorization of A as dcscribcd in [4].

Considcr now rcgularization in standard form. As can bc sccn from Eq. (2), thc additional rcquircmcnt on thc norm of thc solution cntcrs dircctly into thc dcfinition of thc rcgularizcd solution x_λ. For thcorctical invcstigations, this solution can also be cxprcsscd in tcrms of thc SVD of A. 10 do this, it is convcnicnt to writc x_λ as

$$x_\lambda = A_\lambda^I b \tag{10}$$

whcrc thc matrix $A_\lambda^I \in R^{n\times m}$ is a "rcgularizcd invcrsc", dcfincd by:

$$A_\lambda^I \equiv (A^T A + \lambda^2 I,)-' A^T . \tag{11}$$

(A_λ^I is only a {3,4}-invcrsc of A [2] and thcrcforc not rcally an invcrsc.) This matrix turns out to be closcly rclatcd to thc matrix A_k^+ abovc. To scc this, introducc thc matrix

$$\Sigma_\lambda^+ \equiv \mathrm{diag}\left\{ \frac{\sigma_1}{\sigma_1^2 + \lambda^2}, \dots, \frac{\sigma_n}{\sigma_n^2 + \lambda^2} \right\} \in R^{m\times n} . \tag{12}$$

Whcn (4) is insertcd into (11), it is sccn that $A,'$ can bc writtcn in tcrms of thc SVD of A as

$$A_\lambda^I = V \Sigma_\lambda^+ U^T . \tag{13}$$

This cstablishcs thc nicc similaritics bctwccn Eqs. (8), (9) and (IO), (13), rcspcctivcly. Thc matrix A_λ^I should not bc computcd in any of thc forms (11) or (1 3); instcad, x_λ can bc computcd cfficicntly dircctly from (3) as dcscribcd in [6].

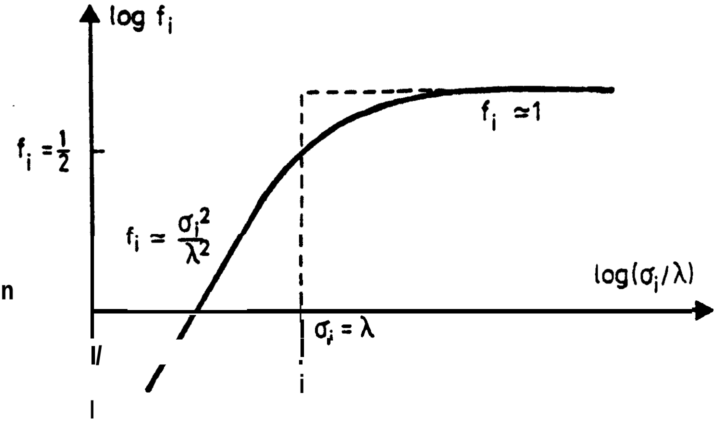
Thc most important obscrvation from Eqs. (9) and (13) is that rcgularization, likc thc TSVD, tcnds to filtcr out thc contributions to thc solution corrcsponding to thc smallcst singular valucs [22]. To claboratc on this, thc ith diagonal clcmcnt of Σ_k^+ as wcll as Σ_λ^+ can bc writtcn as thc ith diagonal clcmcnt of Σ_r^+ timcs a *filtcr factor* f_i. For thc TSVD, this filtcr factor has thc form

$$f_i = \begin{cases} 1 & \text{for } \sigma_i \geq \sigma_k \\ 0 & \text{for } \sigma_i < \sigma_k \end{cases} \tag{14}$$

corrcsponding to a sharp filtcr that simply cuts off thc last $n - k$ componcnts. For rcgularization, thc filtcr factor takcs thc form

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} , \quad i = 1, 2, \dots, n \tag{15}$$

corrcsponding to a smooth filtcr that damps thc componcnts corrcsponding to $\sigma_i < \lambda$. Whcn k is choscn such that $\sigma_k = \lambda$, thc sharp filtcr of thc TSVD can in fact bc sccn as an approximation to thc smooth filtcr of thc rcgularization mcthod, cf. Fig. 1. This can bc takcn as a hint that x_k and x_λ may bc similar, and in Scctions 5 and 6 this will bc invcstigatcd furthcr.



Figurc 1.
--- Filtcr factor for rcgularization in standard form.
- - - Filtcr corrcsponding to truncatcd SVD with $\sigma_k = \lambda$.

3. Perturbation theory for the TSVD

To investigate the circumstances in which the TSVD is applicable as a regularization method, it is necessaty to have a perturbation theory for the TSVD. Such a perturbation analysis is carried out in this section. The results here are strongly connected with those of Wedin [23,24].

$$\tilde{A} = A + E = \tilde{U}\tilde{\Sigma}\tilde{V}^T \quad , \quad \tilde{b} = b + e \tag{16}$$

and let $\tilde{x}_k$ denote the perturbed TSVD solution

$$\tilde{x}_k = \tilde{A}_k^+ \tilde{b} \tag{17}$$

with $\tilde{A}_k$ and $\tilde{A}_k^+$ defined as in (7) and (9). Also, define the following three useful quantities:

$$\kappa_k \equiv \|A\|\,\|A_k^+\| = \sigma_1/\sigma_k \tag{18a}$$

$$\eta_k \equiv \|E\|\,\|A_k^+\| = \|E\|/\sigma_k = \kappa_k \frac{\|E\|}{\|A\|} \tag{18b}$$

$$\omega_k \equiv \|A - A_k\|\,\|A_k^+\| = \sigma_{k+1}/\sigma_k \quad . \tag{18c}$$

κ_k is generally known as *the condition number* of A_k, η_k is equal to κ_k times the relative error level $\|E\|/\|A\|$, andω_k is the size of the *relative gap* in the singular value spectrum between singular values σ_k and σ_{k+1}.

Theorem 3.1. *Assume that* $\|E\| < \sigma_k$. *Then:*

$$\|\tilde{A}_k^+\| = \frac{1}{\tilde{\sigma}_k} \leq \frac{1}{\sigma_k - \|E\|} = \frac{\|A_k^+\|}{1 - \eta_k} \quad . \tag{19}$$

Theorem 3.1 can also be found in e.g. [24, Lemma 3.1], but is included here for completeness. It states that $\|\tilde{A}_k^+\|$ increases monotonically when the norm $\|E\|$ approaches σ_k. Hence, for the TSVD to be useful, $\|E\|$ must be small compared to the k'th singular value of A, otherwise $\tilde{A}_k^+$ may differ considerably from A_k^+. This point is elaborated in the following theorem:

Theorem 3.2. *Assume that* $\|E\| < \sigma_k - \sigma_{k+1}$. *Then the relative error of* $\|A_k^+\|$ *is bounded by:*

$$\frac{\|A_k^+ - \tilde{A}_k^+\|}{\|A_k^+\|} \leq 3\,\frac{\kappa_k}{(1-\eta_k)(1-\eta_k - \omega_k)}\,\frac{\|E\|}{\|A\|} \quad . \tag{20}$$

As a special case, *for* **k** = *r* = rank(A) *and* $\|E\| < \sigma_r$:

$$\frac{\|A^+ - \tilde{A}^+\|}{\|A_+\|} \leq 3\,\frac{\kappa_r}{1-\eta_r}\,\frac{\|E\|}{\|A\|} \quad . \tag{21}$$

Proof. The proof follows from [23, Eq. (4.6)] and Theorem 3.1 above:

$$\begin{aligned}
\|A_k^+ - \tilde{A}_k^+\| &\leq \|E\|\left[\|A_k^+\|\,\|\tilde{A}_k^+\| + \frac{\|A_k^+\| + \|\tilde{A}_k^+\|}{\sigma_k - \tilde{\sigma}_{k+1}}\right] \\
&\leq \|E\|\left[\frac{\|A_k^+\|^2}{1-\eta_k} + \frac{\|A_k^+\| + \|A_k^+\|/(1-\eta_k)}{\sigma_k - \tilde{\sigma}_{k+1}}\right] \\
&= \|A_k^+\|\eta_k\left[\frac{1}{1-\eta_k} + \frac{2-\eta_k}{(1-\eta_k)(1-\tilde{\sigma}_{k+1}/\sigma_k)}\right] \\
&= \|A_k^+\|\eta_k\,\frac{3-\eta_k-\tilde{\sigma}_{k+1}/\sigma_k}{(1-\eta_k)(1-\tilde{\sigma}_{k+1}/\sigma_k)} \leq \frac{3\eta_k\|A_k^+\|}{(1-\eta_k)(1-\eta_k-\omega_k)} \quad .
\end{aligned} \tag{22}$$

Insertion of (18b) in (22) then yields (20). (21) follows from the fact that if $\|E\| < \sigma_r$, then

$\mathrm{rank}(\tilde{A}) = \mathrm{rank}(A)$ and $\tilde{\sigma}_{r+1} = 0$. □

Eq. (21) is a well-known result and states that for $\tilde{A}^+$ to be close to $A+$, both the condition number κ_k and the quantity η_k must be small. Eq. (20) is an extension of this result to the TSVD, and it is seen that for the perturbed pseudoinverse of the TSVD, $\tilde{A}_k^+$, to be close to A_k^+ it is also necessary to require that the relative gap ω_k be small. In other words, if the SVD is to be successfully truncated at k, then there must be a well-determined gap between the singular values σ_k and σ_{k+1}. This is also the essence of the following extension Of [23, Eq. (3.1)].

Theorem 3.3. *Assume that* $\| E \| < \sigma_k - \sigma_{k+1}$ *and let* θ_k *denote the subspace angle*

$$\theta_k = \theta\{S(A_k), S(\tilde{A}_k)\} \tag{23}$$

where S is any of the four fundamental subspaces N, N-, R, and R-. Then:

$$\sin\theta_k \leq \frac{\|E\|}{\sigma_k - \tilde{\sigma}_{k+1}} = \frac{\eta_k}{1 - \tilde{\sigma}_{k+1}/\sigma_k}$$
$$\leq \frac{\eta_k}{1 - \eta_k - \omega_k} = \frac{\kappa_k}{1 - \eta_k - \omega_k} \frac{\|E\|}{\|A\|} . \tag{24}$$

As a special case, for $k = r = \mathrm{rank(A)}$ *and* $\|E\| < \sigma_r$*:*

$$\sin\theta_r \leq \eta_r = \|E\| \, \|A^+\| = \kappa_r \frac{\|E\|}{\|A\|} . \tag{25}$$

Finally, consider the perhaps most important result: the relative perturbation of the TSVD solution (8). The following theorem is an extension of [24, Theorem 5.1].

Theorem 3.4. *Assume that* $\| E \| < \sigma_k - \sigma_{k+1}$, *and, let* $r_k = b - A x_k$ *denote the residual corresponding to the TSVD solution* x_k*. Then:*

$$\frac{\|x_k - \tilde{x}_k\|}{\|x_k\|} \leq \frac{\kappa_k}{1 - \eta_k} \left[\frac{\|E\|}{\|A\|} + \frac{\|e\|}{\|b\|} \quad \frac{\eta_k}{1 - \eta_k - \omega_k} \|r_k\| \right] + \frac{\eta_k}{1 - \eta_k - \omega_k} . \tag{26}$$

As a special case, for $k = r = \mathrm{rank}\,(A)$ *and* $\| E \| < \sigma_r$*:*

$$\frac{\|x_r - \tilde{x}_r\|}{\|x_r\|} \leq \frac{\kappa_r}{1 - \eta_r} \left[\frac{\|E\|}{\|A\|} + \frac{\|e\|}{\|b\|} + \eta_r \frac{\|r_r\|}{\|b\|} \right] + \eta_r . \tag{27}$$

In both equations, the denominator $\| b \|$ *can be replaced by* $\| A x_k$ *II and* $\| A \| \| x_k \|$, *thus tightening the bounds.*

Proof: This proof follows a different line than that of [24]. The error of the TSVD solution is:

$$\tilde{x}_k - x_k = \tilde{A}_k^+ \tilde{b} - x_k = \tilde{A}_k^+ (b + e) - x_k = \tilde{A}_k^+ (A x_k + r_k + e) - x_k$$
$$= \tilde{A}_k^+ ((\tilde{A} - E) x_k + r_k + e) - x_k = \tilde{A}_k^+ (\tilde{A} x_k - E x_k + r_k + e) - x_k$$
$$= \tilde{A}_k^+ (- E x_k + e + r_k) - (I_n - \tilde{A}_k^+ \tilde{A}_k) x_k$$

Taking norms on both sides yields

$$\|\tilde{x}_k - x_k\| \leq \|\tilde{A}_k^+\| \left[\|E\| \, \|x_k\| + \|e\| + \|r_k\| \right] + \|(I_n - \tilde{A}_k^+ \tilde{A}_k) x_k\| \tag{28}$$

The contribution to' $\|\tilde{x}_k - x_k\|$ from the vector $- E x_k + e + r_k$ comes from its component in $R(\tilde{A}_k)$, the range of $\tilde{A}_k$, and in the worst case both $E x_k$ and e belong to $R(\tilde{A}_k)$. The contribution from r_k is, however, bounded by $\| r_k \| \sin \theta_k$, where $\theta_k = 8 \{ R(A_k), R(\tilde{A}_k)\}$, cf. Fig. 2(a). Concerning the second term in (28), the matrix $I_n - \tilde{A}_k^+ \tilde{A}_k$ is the projection matrix for orthogonal projection onto $N(\tilde{A}_k)$, the -null space of $\tilde{A}_k$, and from Fig. 2(b) it follows that $\|(I_n - \tilde{A}_k^+ \tilde{A}_k) x_k\| = \|x_k\| \sin\varphi_k$, where $\varphi_k = \varphi\{N\text{-}(A_k), \tilde{N}\text{-}(A_k)\}$. Upper bounds for both angles θ_k and φ_k are given in Theorem 3.3, and an upper bound for $\| \tilde{A}_k^+ \|$ is given in Theorem 3.1. This gives Eqs. (26) and (27). □

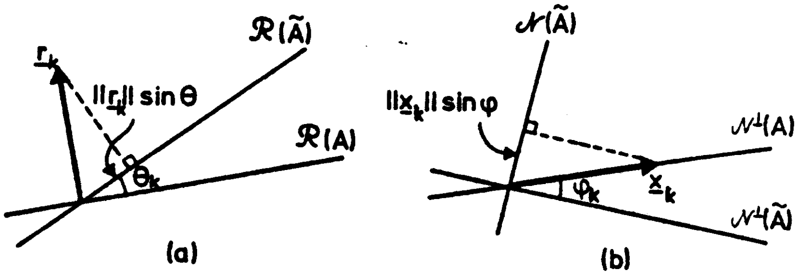


Figure 2. The contributions to (28) from (a) r_k and (b) x_k.

Again, Eq. (27) is a well-known result. The new main result (26) supplements Theorems 3.2-3.3 above and states again that the perturbed TSVD solution x_k can only be guaranteed to be close to the true solution when there is a well-determined gap between singular values σ_k and σ_{k+1}.

4. Matrices with well-determined and ill-determined numerical rank

Although the concept of matrix rank is not necessary for the use of TSVD as a method for regularization, it is appropriate to discuss this concept here since it is so strongly connected with the above perturbation theory. This leads to a natural division of ill-conditioned matrices into two classes: those with well-determined numerical rank and those with ill-determined numerical rank. Such a characterization was also discussed by Golub, Klema & Stewart [9].

It is well-known that, due to approximations, rounding errors, and other sources of errors, it is very unlikely that true zero singular values occur in practical numerical applications. It is therefore common to neglect the singular values smaller than a certain threshold, which obviously corresponds to the use of the TSVD. The choice of a suitable threshold takes its basis in the following classical perturbation bound for the singular values [10, p. 286]:

$$|\sigma_i - \tilde{\sigma}_i| \le \|E\| \quad , \quad i = 1,2,\dots,n \quad . \tag{29}$$

This implies that singular values $\tilde{\sigma}_i$ of $\tilde{A}$ larger than $\|E\|$ are guaranteed to represent nonzero singular values σ_i of A. However, one cannot distinguish the singular values $\tilde{\sigma}_i$ below $\|E\|$ from exact zeroes. As a consequence, when

$$\tilde{\sigma}_k > \|E\| \ge \tilde{\sigma}_{k+1} \tag{30}$$

for some k, one can only guarantee that the rank of A is at least k.

This leads to the definition of the *numerical rank* r_τ of A, with respect to the error level $\tau > 0$, as the number of singular values strictly greater than τ:

$$\sigma_1 \ge \cdots \ge \sigma_{r_\tau} > \tau \ge \sigma_{r_\tau+1} \tag{31}$$

Equivalent names for 'numerical rank' are 'effective rank' [8] and 'pseudorank' [12]. The so defined TSVD A_{r_τ} (7) consists only of those contributions $u_i\, \sigma_i\, v_i^T$ to A with a significant magnitude as measured by the error level τ, while the uncertain contributions (corresponding to $i > r_\tau$) are discarded.

The above .definition of the numerical rank is indcpendent of the particular distribution of the singular values of A. The numerical rank r_τ is, however, only useful when r_τ is well-determined with respect to τ; i.e., r_τ must be insensitive to small variations in τ. As seen from Fig. 3, this is only the case when there is a well-defined gap between the singular values σ_{r_τ} and $\sigma_{r_\tau+1}$. Exactly the same conclusion can be drawn from Theorems 3.2-3.4 of the perturbation theory in the abovesection, with $k = r_\tau$. Depending on the behavior of the singular value spectrum, it is therefore natural to characterize an ill-conditioned matrix as either a matrix with a ***well-determined numerical rank*** or ***an ill-determined numerical rank*** One should look for an numerical rank only if it actually can be expected to be there, as is the case for ill-conditioned matrices with a well-determined numerical rank.

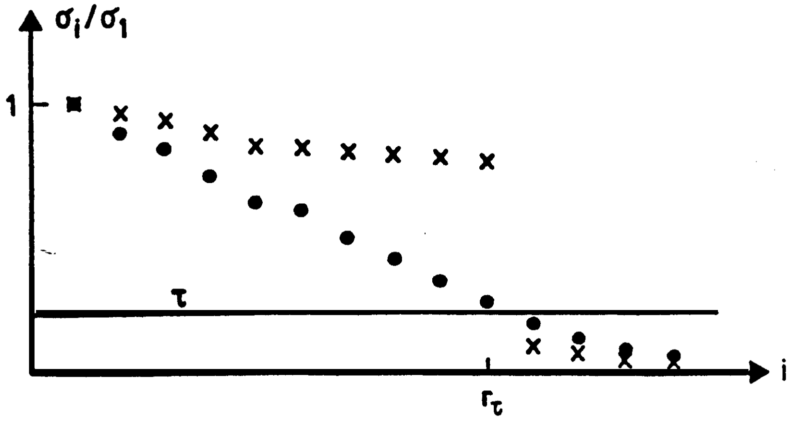


Figure 3. Singular value spectra corresponding to an ill-conditioned matrix with X well-determined and • ill-determined numerical rank.

It should be noted that the scaling of A has a considerable effect on its singular value spec trum. Inherent in the above discussion is therefore that the matrix A has been properly scaled. Good scaling strategies seem to be to scale so that, as far as possible, the uncertainties in all the ele-. ments of A are of the same order of magnitude, or so that ail columns of A have approximately the same norm $\|\,.\,\|$.

Any distribution of singular values in between the two extremes of Fig. 3 may of course be expected in practical applications. However, there are certain entegories of problems that clearly lead to ill-conditioned matrices with either well-determined or ill-determined numerical rank. Matrices with well-determined numerical rank are most likely to occur when the algebraic least squares problem (1) is obtained from some underlying problem for which the concept of rank makes sense. Examples of such problems are:

observation of signal components in noisy data [19],

solution of some Fredholm integral equations of the first kind [7,11],

determination of (A ,B)-invariant and controllability subspaces [15,20].

Matrices with ill-determined numerical rank, on the other hand, are obtained from underlying ill-posed problems where the concept of rank has no intuitive interpretation. Examples of such problems are:

digital image restoration [1],

solution of integral equations in solid state physics [5],

inverse Radon and Laplace transformation [14,16,21].

5. TSVD and matrices with well-determined numerical rank

Whcn the matrix A in problem (1) has a well-dctcrmincd numerical rank as discussed in Scction 4, it sccms natural to apply the TSVD as a mcthod for rcgularization. The question then arises: undcr which circumstances is the TSVD solution x_k closc to the rcgularizcd solution x_λ? This obviously dcpcnds on the choicc of the rcgularization parameter A, and thc filtcr factors (14) and (15) in Scction 2 suggest that A should bc chosen somcwhcrc bctwccn the singular valucs σ_k and σ_{k+1}, wherc k is the numcrical rank. The choicc of thc regularization parameter A is not a trivial problem as can bc seen from the following theorem.

Thcorcm 5.1. ***Let* λ *be chosen such that* $\lambda \in [\sigma_{k+1}, \sigma_k]$. *Then:***

$$\frac{\|A_k^+\|}{2} \le \|A_\lambda^I\| \le \frac{\|A_k^+\|}{2\,\omega_k} \tag{32}$$

where* ω_k *is the relative gap (18c).

Proof. From Eq. (13) is follows that

$$\|A_\lambda^I\| \equiv \max_i \left\{ \frac{\sigma_i}{\sigma_i^2 + \lambda^2} \right\} = \max \left\{ \frac{\sigma_k}{\sigma_k^2 + \lambda^2} , \frac{\sigma_{k+1}}{\sigma_{k+1}^2 + \lambda^2} \right\} \quad .$$

Both of thcse quantitics arc monotonically decreasing functions of λ, and they intersect at $\lambda = (\sigma_k\ \sigma_{k+1})^{1/2}$. Thus, the minimum and maximum values of $\| A_\lambda^I$ II arc attained at $\lambda = \sigma_k$ and $\lambda = \sigma_{k+1}$, rcspcctively:

$$\lambda = \sigma_k \quad => \|A_\lambda^I\| = \tfrac{1}{2}\sigma_k^{-1} = \tfrac{1}{2}\|A_k^+\| \quad ,$$

$$\lambda = \sigma_{k+1} => \|A_\lambda^I\| = \tfrac{1}{2}\sigma_{k+1}^{-1} = \tfrac{1}{2}\sigma_k^{-1}\omega_k^{-} = \tfrac{1}{2}\|A_k^+\|\omega_k^{-1} \ . \qquad \square$$

Theorem 5.1 states that if ω_k is small, i.e. if therc is a large gap in the singular valuc spectrum of A at k, thcn A_λ^I may differ considcrably from A_k^+ if λ is chosen closc to σ_{k+1}. On the other hand, if λ is chosen closc to σ_k, then A_λ^I and A_k^+ i-night bc similar. The closcncss of thcsc two matriccs is invcstigatcd in the following theorem.

Thcorcm 5.2. ***Assume that* $\lambda \in [\sigma_{k+1}, \sigma_k]$. *Then:***

$$\frac{\omega_k^{1/2}}{1+\omega_k^{1/2}} \le \min_\lambda \frac{\|A_\lambda^I - A_k^+\|}{\|A_k^+\|} \le \frac{\omega_k^{1/2}}{1+\omega_k^{3/2}} \tag{33}$$

and the minimum is obtained when

$$\lambda \approx (\sigma_k^3\,\sigma_{k+1})^{1/4} \ . \tag{33a}$$

Under the same assumption:

$$\min_\lambda \|A(A_\lambda^I - A_k^+)\| = \frac{\omega_k}{1+\omega_k} \tag{34}$$

and this minimum is obtained for

$$\lambda = (\sigma_k\ \sigma_{k+1})^{1/2} \quad . \tag{34a}$$

Proof. Thc proof follows dircctly from thc hppcndix and thc following rclations:

$$\|A_\lambda^I - A_k^+\| = \|\Sigma_\lambda^+ - \Sigma_k^+\| = \max_i |\,[\xi_\lambda - \xi_k]_i\,| \quad \text{for } p = 0 \ ,$$

$$\|A(A_\lambda^I - A_k^+)\| = \|\Sigma(\Sigma_\lambda^+ - \Sigma_k^+)\| = \max_i |\,[\xi_\lambda - \xi_k]_i\,| \quad \text{for } p = 1 \ . \qquad \square$$

Thc csscncc of Theorem 5.2 is that the diffcrcncc between thc matrices A^I_λ and A_k^+ is small whcn λ is chosen according to (33a) ***and*** when ω_k is small, in which casc the minimum (33) is approximatcly cqual to $\omega_k^{\frac{1}{2}}$. Thc difference bctween thc solutions x_λ and x_k is thercfore also small for this valuc of λ. Similarly, it is sccn that- the diffcrcnce bctwccn the residuals corrcsponding to the two solution:

$$(b - A\, x_k) - (b - A\, x_\lambda) = A\,(A^I_\lambda - A_k^+)\,b \ , \tag{35}$$

is small when λ is chosen according to (34a) *and* when ω_k is small, in which casc the minimum (34) is approximatcly equal to ω_k. There is a trade-off bctwcen (33) and (34); but since (34) is a factor $\omega_k^{\frac{1}{2}}$ smallcr than (33) is seems appropriatc to choosc λ according to (33a).

To illustrate the impact of the above thcorem it is convenicnt to compare thc two methods graphically as in Fig. 4. The solid curve, which is associated with the regularization method, is given by

$$(\|b_R - A\, x_\lambda\|, \|x_\lambda\|) \ , \quad \lambda \geq 0 \tag{36}$$

whcre b_R is the projcction of the right-hand sidc on the range of A. Similarly, thc points marked X are associatcd with thc TSVD and reprcsent the point set

$$(\|b_R - A\ x_k\|, \|x_k\|) \ , \quad k = 0, 1, ..., n \quad . \tag{37}$$

The bchavior of (36) and (37) is described in the following thcorcm.

Theorem 5.3. *In (36). $\| x_\lambda \|$ is a decreasing function of $\|\, b_R - A\, x_\lambda \|$, and in (37), $\| x_k \|$ is a decreasing function of $\| b_R - A\, x_k \|$ on a finite set. The curve coincides with the point set at the endpoints $(k = 0, \lambda = 0)$ and $(k = n, \lambda = \infty)$ where they both touch the axes. The remaining points of (37) lie above the solid curve (36).*

Proof The fact that $\| x_\lambda \|$ and $\| x_k \|$ arc dccreasing functions follows. from Eqs. (3) and (10) and the following exprcssions:

$$\|b_R - A\ x_\lambda\| = \sum_{i=1}^{n} \frac{\lambda^2}{\sigma_i^2 + \lambda^2} (u_i^T b)^2 \ , \quad \|b_R - A\ x_k\| = \sum_{i=k+1}^{n} (u_i^T b)^2$$

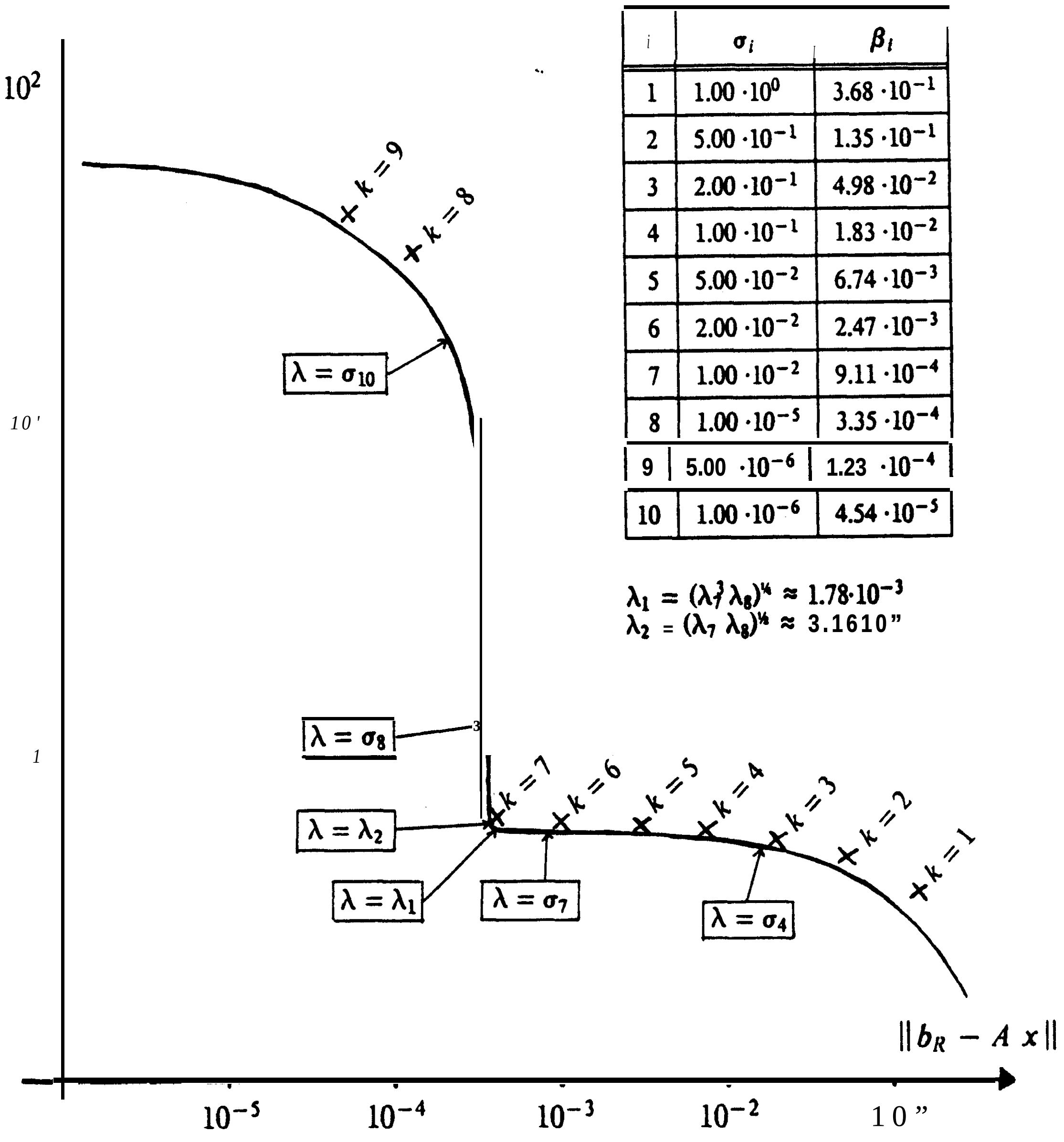
in which u_i is. the i'th column of thc matrix U in (4). Eldén [6] has shown that x_λ can also bc charactcrizcd by

$$\min \|b - A\, x\| \text{ subjcct t o } \|x\| \leq 7$$

whcre 7 is a frcc paramctcr, and that normally thc solution occurs whcn $\|x\| = 7$. Hcncc, $\|x\| \geq \|x_\lambda\|$ for any x that satisfics $\| b_R - A\ x \| = \| b_R - A\ x_\lambda \|$. ❑

Fig. 4 is drawn for thc casc whcn A has a well-dctcrmincd numcrical rank, and thc 'corncr' of thc solid curvc is charactcristic for such matriccs. It is intuitivcly clcar that in ordcr to yicld a fair tradc-off bctwccn minimization of thc rcsidual norm and thc solution norm (2), λ should bc chosen such that x_λ is rcprcscntcd by a point ncar thc 'corncr' of thc curvc. Thc figurc shows that this is actually thc casc whcn λ is chosen according to (33a) as wcll as (34a). Thc figurc also shows that thc TSVD solution x_k is in fact closc to x_λ for this choicc of λ.

Thc conclusion to bc drawn from this discussion is that if thc matrix A is ill-conditioned and has a wcll-dctcrmincd numcrical rank, and if λ is chosen ncar thc intuitivc optimum valuc, thcn thc TSVD solution x_k is guarantccd to bc similar to thc rcguiarizcd solution x_λ. This suggcsts that for this class of matriccs, from a thcorctical as wcll as a computational point of vicw, a suitablc solution to (1) is thc TSVD solution x_k which, in most cascs, can bc computcd cfficicntly from a Q-R factorization of thc matrix A [4].



i	σ_i	β_i
1	$1.00 \cdot 10^{0}$	$3.68 \cdot 10^{-1}$
2	$5.00 \cdot 10^{-1}$	$1.35 \cdot 10^{-1}$
3	$2.00 \cdot 10^{-1}$	$4.98 \cdot 10^{-2}$
4	$1.00 \cdot 10^{-1}$	$1.83 \cdot 10^{-2}$
5	$5.00 \cdot 10^{-2}$	$6.74 \cdot 10^{-3}$
6	$2.00 \cdot 10^{-2}$	$2.47 \cdot 10^{-3}$
7	$1.00 \cdot 10^{-2}$	$9.11 \cdot 10^{-4}$
8	$1.00 \cdot 10^{-5}$	$3.35 \cdot 10^{-4}$
9	$5.00 \cdot 10^{-6}$	$1.23 \cdot 10^{-4}$
10	$1.00 \cdot 10^{-6}$	$4.54 \cdot 10^{-5}$

$\lambda_1 = (\lambda_7^3 \lambda_8)^{1/4} \approx 1.78 \cdot 10^{-3}$
$\lambda_2 = (\lambda_7 \lambda_8)^{1/2} \approx 3.1610”$

Figure 4. Comparison of the TSVD and regularization methods for an ill-conditioned matrix with well-determined numerical rank. The numerical rank is obviously equal to $k = 7$.

6. TSVD and matrices with ill-determined numerical rank

In spite of the conclusion from the previous section, the TSVD has also been reported [22] to yield results similar to those of regularization when the matrix A has an ill-determined numerical rank. For such matrices, Theorem 5.2 is not useful since it turns out that the similarity of x_λ and x_k now depends strongly on the right-hand side b. More precisely, it depends on the projection of b on the left singular vectors of A, as can be seen from the following theorem.

Theorem 6.1. *Let $\beta_i = u_i^T b$, $i = 1, \ldots, n$, where $\{u_i\}$ are the columns of the left singular matrix U in the SVD of A. Assume that the $\{\beta_i\}$ decay as:*

$$\beta_i = \sigma_i^p \quad , \; p = 0, 1, 2, 3, 4 \tag{38}$$

and assume that λ is chosen such that $\lambda \in [\sigma_{k+1}, \sigma_k]$. Then the difference between x_λ and x_k as measured by

$$d \equiv \min_\lambda \|x_\lambda - x_k\|_\infty \quad , \quad \lambda \in [\sigma_{k+1}, \sigma_k] \tag{39}$$

is a function of ω_k and κ_k as shown in Table 1 below. When $\omega_k \approx 1$, which corresponds to an A with ill-determined numerical rank, the relative measure $d \|b\|_\infty^{-1}$ is a function of κ_k only as shown in Table 1.

P	β_i	$d.$	$\frac{d}{\|b\|_\infty}, \omega_k \approx 1$
0	1	$\frac{1}{\sigma_k} \frac{\omega_k^{1/2}}{1+\omega_k^{1/2}} \le d \le \frac{1}{\sigma_k} \frac{\omega_k^{1/2}}{1+\omega_k^{3/2}}$	$\sigma_1^{-1} \frac{\kappa_k}{2}$
1	σ_i	$\frac{\omega_k}{1+\omega_k}$	$\sigma_1^{-1} \frac{1}{2}$
2	σ_i^2	$\sigma_k \frac{\omega_k^{3/2}}{1+\omega_k^{1/2}} \le d \le \sigma_k \frac{\omega_k^{3/2}}{1+\omega_k^{3/2}}$	$\sigma_1^{-1} \frac{1}{2\kappa_k}$
3	σ_i^3	$\sigma_1^2 \frac{\omega_k^2}{\kappa_k^2+\omega_k^2}$	$\sigma_1^{-1} \frac{1}{\kappa_k^2}$
4	σ_i^4	$\sigma_1^3 \frac{\omega_k^2}{\kappa_k^2+\omega_k^2}$	$\sigma_1^{-1} \frac{1}{\kappa_k^2}$

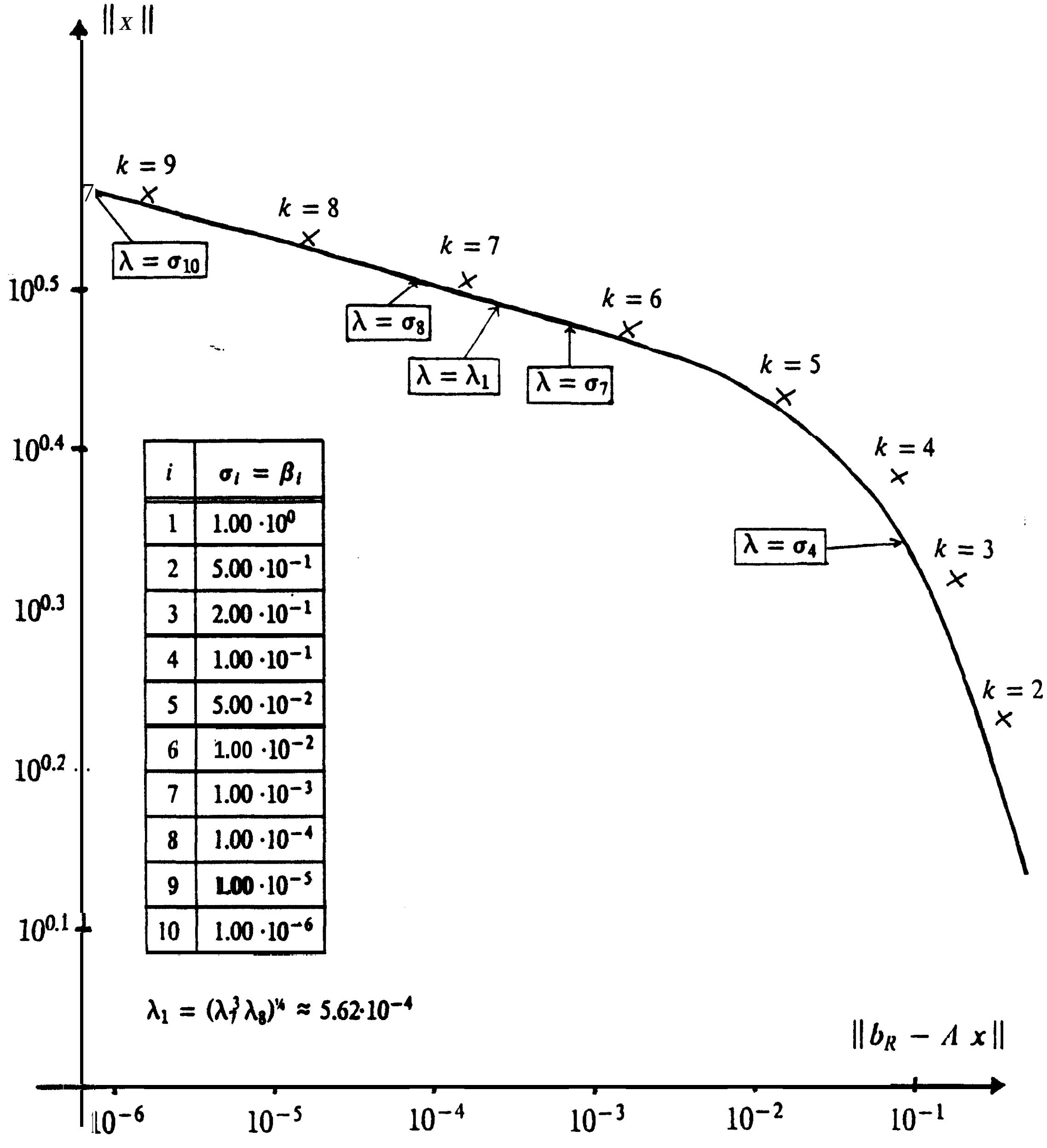
Table 1. The similarity of x_λ and x_k as a function the decay of $\{\beta_i\}$.

Proof. See the Appendix.

Although Theorem 6.1 covers only a very special case of right-hand sides b it gives a clear indication of the importance of the decay of the $\{\beta_i\}$. For the case of Fredholm integral equations of the first kind. the well-known Picard condition (cf. e.g. [22]) states that for a solution to exist, the corresponding β_i-coefficients must decay faster than the singular values $\{\mu_i\}$ of the kernel such that $\Sigma (\beta_i / \mu_i)^2 < \infty$. Theorem 6.1 is a kind of 'discrete Picard condition' for the TSVD and states that the faster decay of the $\{\beta_i\}$, the closer x_k gets to x_λ.

The case $p = 0$ is unrealistic for most practical right-hand sides b and is included here only for completeness. The case $p = 1$ is slightly unrealistic; but it sometimes occurs in practical applications. Both cases do, however, apply to the perturbation e of b in many practical applications when the right-hand side b consists of measured quantities contaminnted with measurement errors.

If, on thc othcr hand, the cocfficients $\{u_i^T e\}$ of thc perturbation e decay like the β_i-cocfficients for $p \geq 2$, thcn Thcorcm 6.1 lcads to a small perturbation bound on thc TSVD solution. This result agrees with the pcrturbation bound given in terms of the 'effcctivc condition number' as dcfincd in [3].



i	$\sigma_i = \beta_i$
1	$1.00 \cdot 10^0$
2	$5.00 \cdot 10^{-1}$
3	$2.00 \cdot 10^{-1}$
4	$1.00 \cdot 10^{-1}$
5	$5.00 \cdot 10^{-2}$
6	$1.00 \cdot 10^{-2}$
7	$1.00 \cdot 10^{-3}$
8	$1.00 \cdot 10^{-4}$
9	$1.00 \cdot 10^{-5}$
10	$1.00 \cdot 10^{-6}$

$\lambda_1 = (\lambda_7^3 \lambda_8)^{1/4} \approx 5.62 \cdot 10^{-4}$

Figurc 5. Comparison of the TSVD and rcgularization methods for an ill-conditioned matrix with ill-dctcrmincd numerical rank. Thcrc is no intuitivc way of choosing λ or *k*.

It is interesting to notice the different behavior of d and $d \, \| \, b \, \|_\infty^{-1}$ for $p \leq 2$ and for $p > 2$. The reason for this behavior is that for $p \leq 2$ the maximum element of $x_\lambda - x_k$ is associated with σ_k, while for $p > 2$ it is associated with σ_1. It is also interesting to note that $d \, \| \, b \, \|_\infty^{-1}$ actually decreases with increasing κ_k, thus suggesting that k be chosen as large as possible to maximize κ_k. The perturbation theory of Section 3 does, however, still apply and there is therefore a trade-off on κ_k between minimization of $d \, \| \, b \, \|_\infty^{-1}$ and minimization of the condition number κ_k.

The general situation is illustrated in Fig. 5, showing the typical behavior of (36) and (37) when A has an ill-determined numerical rank for the case $p = 1$ above. Here, there is obviously no intuitive way of choosing a suitable regularization parameter λ. Neither does the singular value spectrum of A suggest a suitable value of k. x_k will be close to x_λ for any k provided that the 'discrete Picard condition' is satisfied

The conclusion to be drawn in this section is that if A is ill-conditioned and has an ill-: determined numerical rank then the TSVD solution x_k will be close to the regularized solution x_λ if the β_i-coefficients of the right-hand side b decay sufficiently fast, Hence, use of the TSVD as a regularization method might give good results. In general, one can not guarantee a small perturbation bound on the solution x_k for any value of k; but if the perturbation e of the right-hand side also satisfies the 'discrete Picard condition' then the perturbation bound on x_k is small.

7. Conclusion

From a theoretical as well as a practical point of view, the truncated singular value decomposition (TSVD) is a suitable method for regularization of the ill-posed problem (1) when the coefficient matrix. A is ill-conditioned with a well-determined numerical rank. If the parameter k of the TSVD A_k (7) is chosen equal to the numerical rank r_τ (31) of A, then the TSVD solution is little sensitive to errors in the matrix A and right-hand side b, and the TSVD solution is close to the regularized solution with the regularization parameter chosen near its intuitive optimum value. When A has an ill-determined numerical rank, the TSVD and regularization methods may also produce similar results, provided that the β_i-coefficients (38) of b decay sufficiently fast., and if the. corresponding coefficients of the perturbation e of b also decay sufficiently fast then the TSVD solution is little sensitive to these errors.

Acknowledgements

I would like to thank Prof. Gene H. Golub for the invitation to visit Stanford University and for providing nice working conditions. Prof. Åke Björck and Prof. Hans Bruun Nielsen are acknowledged for valuable discussions and suggestions. I am indebted to Prof. Per-Åke Wedin for his great influence on this presentation.

Appendix: Proofs of Theorems 5.2 and 6.1

In this appendix, the quantity d as defined in Eq. (39) is investigated for the special case when $\beta_i = \sigma_i{}^p$ as in Eq. (38). Write $x_\lambda = V \, \xi_\lambda$ and $x_k = V \, \xi_k$, where V is the right singular matrix in the SVD (4) of A. The elements of ξ_λ and ξ_k are then:

$$[\xi_\lambda]_i = \frac{\sigma_i{}^{p+1}}{\sigma_i^2 + \lambda^2} \; , \; i = 1, \ldots, n \qquad \text{and} \qquad [\xi_k]_i = \begin{cases} \sigma_i{}^{p-1} \; , & i = 1, \ldots, k \\ 0 \; , & i = k+1, \ldots, n \end{cases} \qquad \text{(A1)}$$

giving

$$[\xi_\lambda - \xi_k]_i = \begin{cases} -\ \sigma_i{}^{p-1}\dfrac{\lambda^2}{\sigma_i^2 + \lambda^2} & , \mathrm{i} = 1, \ldots, k \\ \sigma_i{}^{p-1}\dfrac{\sigma_i^2}{\sigma_i^2 + \lambda^2} & , \quad i = k+1, \ldots, n \end{cases} \tag{A2}$$

and since $\| x \| = \| V\, \xi \| = \| \xi \|$, it follows that:

$$\begin{aligned} d &\equiv \min_\lambda \| x_\lambda - x_k \|_\infty = \min_\lambda \| \xi_\lambda - \xi_k \|_\infty = \min_\lambda \{ \max_i | [\xi_\lambda - \xi_k]_i | \ \} \\ &= \min_\lambda \left\{ \max_i \left\{ \frac{\sigma_1{}^{p-1}\lambda^2}{\sigma_1^2 + \lambda^2}, \ldots, \frac{\sigma_k{}^{p-1}\lambda^2}{\sigma_k^2 + \lambda^2}, \frac{\sigma_{k+1}^{p+1}}{\sigma_{k+1}^2 + \lambda^2}, \ldots, \frac{\sigma_n^{p+1}}{\sigma_n^2 + \lambda^2} \right\} \right\} \quad . \end{aligned} \tag{A3}$$

Consider first the situation when $\lambda = \sigma_k$. In this case, it is straightforward (but quite cumbersome) to show that:

$$d = f(\lambda, p) = \begin{cases} \dfrac{\sigma_k{}^{p-1}\lambda^2}{\sigma_k^2 + \lambda^2} = \tfrac{1}{2}\sigma_k{}^{p-1} & , \ p = 0,\ 1, 2 \\ \dfrac{\sigma_1{}^{p-1}\lambda^2}{\sigma_1^2 + \lambda^2} = \dfrac{\sigma_1{}^{p-1}\sigma_k^2}{\sigma_1^2 + \sigma_k^2} & , \ p = 3, 4 \end{cases} \quad . \tag{A4}$$

Similarly, when $\lambda = \sigma_{k+1}$ one can show that:

$$d = g(\lambda, p) = \begin{cases} \dfrac{\sigma_{k+1}^{p+1}}{\sigma_{k+1}^2 + \lambda^2} = \tfrac{1}{2}\sigma_{k+1}^{p-1} & , \ p = 0, 1, 2 \\ \dfrac{\sigma_1{}^{p-1}\lambda^2}{\sigma_1^2 + \lambda^2} = \dfrac{\sigma_1{}^{p-1}\ \sigma_{k+1}^2}{\sigma_1^2 + \sigma_{k+1}^2} & , \ \mathrm{P} = 3, 4 \end{cases} \quad . \tag{A5}$$

Hence, for $p = 0, 1, 2$ the maximum element of $\xi_\lambda - \xi_k$ is element no. k or $k+1$, while for $p = 3, 4$ the maximum element is no. 1.

For $p = 0, 1, 2$ the function f is an increasing function of λ, and g is a decreasing function of λ. Hence. d can be written as:

$$d = \min_\lambda \{ f(\lambda, p), g(\lambda, p) \} \tag{A6}$$

and the minimum occurs when $f(\lambda, p) = g(\lambda, p)$; i.e., when elements no. k and $k+1$ are equal. This leads to the equation

$$\frac{\sigma_k{}^{p-1}\lambda^2}{\sigma_k^2 + \lambda^2} = \frac{\sigma_{k+1}^{p+1}}{\sigma_{k+1}^2 + \lambda^2} \tag{A7}$$

which has the solution

$$\frac{\lambda^2}{\sigma_k^2} = \begin{cases} \omega_k^{1/2} \left[-\tfrac{1}{2}\omega_k^{1/2}(\omega_k - 1) + \sqrt{1 + [\tfrac{1}{2}\omega_k^{1/2}(\omega_k - 1)]^2}\ \right] & P = 0 \\ \omega_k & P = 1 \\ \omega_k^{3/2} \left[-\tfrac{1}{2}\omega_k^{1/2}(1 - \omega_k) + \sqrt{1 + [\tfrac{1}{2}\omega_k^{1/2}(1 - \omega_k)]^2}\ \right] & p = 2 \end{cases} \quad . \tag{A8}$$

Insertion of the solution for $p = 1$ into $f(\lambda, 1)$ gives $\omega_k / (1 + \omega_k)$ as given in Table 1. For $\omega_k \ll 1$ and $p = 0, 2$ the above expressions simplify to

$$\frac{\lambda^2}{\sigma_k^2} \approx \begin{cases} \omega_k^{1/2} & , \quad p = 0 \\ \omega_k^{3/2} & , \quad p = 2 \end{cases} \tag{A9}$$

and insertion of this into $f(\lambda, p)$ and $g(\lambda, p)$ gives the upper and lower bounds in the third column of Table 1. For $p = 3, 4$, d can bc written as:

$$d = \min_{\lambda} f(\lambda, p) \tag{A10}$$

and since f is an increasing function of λ, the minimum of (A10) is obtained for $\lambda = \sigma_{k+1}$. This leads to the remaining results in the third column of Table 1. The rightmost column of Table 1 follows from the simple fact that $\| b \|_\infty = \sigma_1$ for all p.

References

1 H.C. Andrews & B.R. Hunt, *Digital image restomtion,* Prentice-Hall (1977).

2 A. Ben-Israel & T.N.E. Grcville, *Genemlized inverses: theory and applications,* Wilcy-Interscience (1974).

3 T.F. Chan & D. Foulser, *Effective condition numbers for linear systems,* Tech Mcmo 86-05, Saxpy Computcr Corporation, Sunnyvale (1986).

4 T.F. Chan & P.C. Hansen, *Truncated SVD solutions by means of rank revealing QR-factorization&* in preparation.

5 U. Eckhardt & K. Mika, *Numerical treatment of incorrectly posed problems - a case study;* in J. Albrecht & L. Collatz *(Eds.), Numerical treatment of integml equations Workshop on numerical treatment of integml equations Oberwolfach. November 18-24,* 1979, Birkhäuser Vcrlag (1980), pp. 92-101.

6 L. Eldén, *Algorithms for regularization of ill-conditioned least squares problems* BIT 17 (1977), 134-145.

7 L. Eldén, *The numerical soluiion of a non-characteristic cauchy problem for a pambolic equation:* in P. Dcuflhard & E. Haircr (Eds.), *Numerical treatment of inverse problems in differential and integral equations* Birkhäuscr Vcrlag (1983); pp. 246-268.

8 G.E. Forsythc, M.A. Malcolm & C.B. Molcr, *Computer methods for mathematical computations,* Prcn ticc-Hall (1977).

9 G.H. Golub, V. Klcma & G.W. Stewart, *Rank degenemcy and least squares problems,* Technical Rcport TK-456, Computcr Scicncc Dcpartmcnt, Univcrsity of Maryland (1976).

10 G.H. Golub & C.F. Van Loan, *Matrix Computations North* Oxford Acadcmic (1983).

11 P.C. Hanscn & S. Christianscn, *An SVD analysis of linear algebraic equations derived from first kind integral equations,* J. Comp. Appl. Math. 1 2&13 (1985), 341-357.

I2 R.J. Hanson, *A numerical method for solving Fredholm integral equations of the first kind using singular values* SIAM J. Numcr. Anal. 8 (1971), 616-622.

13 C.L. Lawson & R.J. Hanson, *Solving least squares problems* Prcnticc Hall (1974).

14 A.K. Louis & F. Nattcrcr, *Mathematical problems of computerized tomogmphy,* Proc. IEEE 71 (1983). 379-389.

15 B.C. Moorc & A.J. Laub, *Computation of supremal (A,B)-invariant and controllability sub spaces* IEEE Trans. Automat. Contr. AC-23 (1978), 783-792.

16 F. Nattcrcr, *Numerical inversion of the Radon transform,* Numcr. Math. 30 (1978), 81-91.

17 D.L. Phillips, *A tcchnique for the numerical solution of certain integral equations of the first kind* J. ACM 9 (1962). 84-97.

18 A.N. Tikhonov, ***Solution of incorrectly formulated problems and the regularization method,*** Dokl. Akad. Nauk. SSSR **151 (1963), 501-504** = Soviet Math. Dokl. 4 **(1963),** 1035-1038.

19 D.W. **Tufts &** R. **Kumaresan,** ***Singular value decomposition and imptvved frequency estimation using linear prediction,*** IEEE Trans. **Acoust.,** Speech, Signal Processing **ASSP-30 (1982), 671-** 675.

20 P.M. Van **Dooren,** ***The generalized eigenstructure problem in linear system theory,*** **IEEE** Trans. Automat. **Contr.** AC-26 (1981) **111-129.**

21 J.M. **Varah,** ***On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems,*** SIAM **J.** Numcr. Anal. 10 **(1973),** 257-267.

22 J.M. Varah, ***A practical examination of some numerical methods for linear discrete ill-posed problems*** SIAM Review 21 **(1979),** 100-111.

23 P.-A. **Wedin,** ***Perturbation bounds in connection with the singular value decomposition,*** BIT 12 **(1972), 99-111.**

24 **P.-Å. Wedin,** ***Perturbation theory for pseudo-inverses,*** **BIT *13* (1973), 217-232.**

25 **P.-Å. Wedin,** ***On the almost rank deficient case of the least squares problem,*** **BIT *13* (1973),** 344-354.