UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Truncated SVD Solutions to Discrete Ill-Posed Problems with Ill-Determined Numerical Rank

Per Christian Hansen

February 1989

CAM Report 89-04

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

Truncated SVD solutions to discrete ill-posed problems with ill-determined numerical rank

Per Christian Hansen

Dept. of Mathematics, UCLA, Los Angeles, CA 90024†

February 1989

ABSTRACT

Tikhonov regularization is a standard method for obtaining smooth solutions to discrete ill-posed problems. A more recent method, based on the singular value decomposition (SVD), is the truncated SVD method. The purpose of this paper is to show, under mild conditions, that the success of both truncated SVD and Tikhonov regularization depends on satisfaction of a discrete Picard condition, involving both the matrix and the right-hand side. When this condition is satisfied, then both methods are guaranteed to produce smooth solutions which are very similar.

AMS subject classification: 65F20, 65R20.

Keywords: Ill-posed problems, truncated SVD, regularization in standard form, perturbation theory.

Abbreviated title: Truncated SVD solutions.

---

† The author is on leave from UNI•C Lyngby, Building 305, Technical University of Denmark, DK-2800 Lyngby, Denmark. He is supported by the Danish Natural Science Foundation, by the National Science Foundation under contract NSF-DMS87-14612, and by the Army Research Office under contract No. DAAL03-88-K-0085.

1. Introduction

This paper is concerned with the linear least squares problem:

$$\min \|A x - b\|_2 , \quad A \in \mathbb{R}^{m \times n} , \quad m \geq n \tag{1}$$

where the matrix A is ill-conditioned and has *ill-determined numerical rank* (i.e., its singular values decay gradually towards zero without any particular gap in the spectrum). Such problems typically arise in connection with the numerical solution of Fredholm integral equations of the first kind,

$$\int K(s,x) f(x)\, dx = g(s) , \tag{2}$$

which are classical examples of ill-posed problems. The following discussion is particularly focused on ill-conditioned least squares problems arising from discretization of (2), but we stress that the results hold for discrete ill-posed problems in general. Throughout the paper, we shall assume for simplicity that the matrix A has full rank.

The singular value decomposition (SVD) is an invaluable tool for analysis of problems with ill-conditioned matrices, and the truncated SVD (TSVD) method has been used successfully to solve a variety of discrete ill-posed problems of the form (1). In spite of this, the method still lacks some theoretical background. Our aim here is to develop a theory for the TSVD and thus provide insight into its behavior. To do this, we find it useful to compare the method with another widely used method for ill-posed problems, namely Tikhonov regularization. This method is theoretically better understood than the TSVD, but there seems to be no general criteria by which these methods can be compared [14]. In [9], Hansen showed that if there is a distinct gap in the singular value spectrum, then TSVD is equivalent to Tikhonov regularization. The present work continues this investigation, with attention primarily focused on matrices with ill-determined numerical rank. We show that the existence of a satisfactory approximate solution primarily depends on satisfaction of a discrete Picard condition and in fact (even for matrices with well-determined numerical rank) has little to do with finding the numerical rank of a matrix. We also show that once the discrete Picard condition is satisfied, then TSVD and Tikhonov regularization always yield very similar solutions. The work was inspired by the 'trilogy' of papers by Varah [22,23,24] and the paper by Aulick & Gallie [1].

The paper is organized as follows. In Section 2 we introduce the methods of truncated SVD and Tikhonov regularization. In Section 3 we show that the convergence of both methods largely depends on the behavior of the right-hand side, and we formulate the discrete Picard condition. Section 4 presents perturbation bounds for the methods, and in Section 5 we further characterize the behavior of the solutions under the influence of errors. Finally, in Section 6 we give two numerical examples.

2. Truncated SVD and Tikhonov regularization

Our investigation takes its basis in the *singular value expansion* (SVE) which is a mean convergent expansion of the kernel K in the form

$$K(s,x) = \sum_{i=1}^{\infty} \mu_i \; u_i(s) \; v_i(x)$$

where both $\{u_i\}$ and $\{v_i\}$ are sequences of orthonormal functions, and all $\mu_i \geq 0$. In terms of the SVE, the solution f to (2) can be written as

$$f(x) = \sum_{i=1}^{\infty} \frac{(u_i\, , g)}{\mu_i} \; v_i(x) \tag{3}$$

where $(u_i\, , g)$ denotes the usual inner product. See e.g. [8, Section 1.2] for more details. The ill-posed nature of (2) is reflected in the facts that the sequence $\{\mu_i\}$ has zero as its only limit point, and the 'smoother' the kernel the faster the μ_i decay to zero [5, Theorem 3.2]. Hence, a square integrable solution f can only exist if the coefficients $(u_i\, , g)$ decay to zero faster than the μ_i, such that

$$\sum_{i=1}^{\infty} \left[\frac{(u_i\, , g)}{\mu_i} \right]^2 < \infty \, . \tag{4}$$

This is the well-known *Picard condition* [8, Theorem 1.2.6].

Corresponding to the SVE of K, the matrix A has a *singular value decomposition* (SVD) in the form [2, Section 3]

$$A = U \, \Sigma \, V^T = \sum_{i=1}^{n} \sigma_i \; \boldsymbol{u}_i \; \boldsymbol{v}_i^T \, ,$$

where the left and right singular vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the orthonormal columns of the matrices $U \in \mathbb{R}^{m \times n}$ and $V \in \mathbb{R}^{n \times n}$, and the singular values σ_i are the diagonal elements of $\Sigma \in \mathbb{R}^{n \times n}$. They satisfy $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ and, since A is assumed to have full rank, $\sigma_n > 0$. The relationship between the ill-posedness of (2) and the large condition number σ_1 / σ_n of the matrix A in (1) was studied by Richter [18] and Wing [25]. Recently, Hansen [10] elaborated on this by showing that whenever (2) is discretized by an expansion method with orthonormal basis functions, the SVD of A is closely related to the SVE of K in the sense that the σ_i and $\boldsymbol{u}_i^T b$ are approximations to the μ_i and

(u_i, g), respectively. This means that if g satisfies the Picard condition (4) and is unaffected by errors, and if the order n of the discretization is sufficiently large, then the *exact least-squares solution* x_o to the unperturbed discretized problem (1), given by

$$x_o \equiv A^+ b = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i \ , \tag{5}$$

yields an approximation to the solution f given in Eq. (3). See also [8, Theorem 4.1.6].

When g and, equivalently, b are perturbed by errors, then the solution to the perturbed problem is very likely to be dominated by errors which are 'blown up' by the small singular values μ_i or σ_i in the denominators of Eqs. (3) or (5). It is therefore necessary to apply some sort of regularization to either (2) or (1) to compute a solution which is less sensitive to the perturbations and which approximates the exact solution to the unperturbed problem, f (3) or x_o (5). Due to the close relationship between the SVD and the SVE, applying a certain regularization method to (1) is equivalent to applying the same regularization method to (2) [10], and the convergence of the regularized algebraic solutions to (1) carries over to the corresponding approximate solutions to (2) [8, Section 4.2].

A highly regarded regularization method, due to Tikhonov [21], amounts to defining the *regularized solution* x_λ as the unique solution to the following least squares problem with a quadratic constraint:

$$\min \{ \|A x - b\|_2^2 + \lambda^2 \|x\|_2^2 \} \ . \tag{6}$$

Here, the *regularization parameter* λ controls the 'smoothness' of the regularized solution. We remind that x_λ can always be written in terms of the SVD as

$$x_\lambda = \sum_{i=1}^{n} \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \frac{u_i^T b}{\sigma_i} v_i \ . \tag{7}$$

If we compare this equation with Eq. (5), we see that the role of the regularization parameter λ is to dampen or filter the terms in the sum corresponding to singular values smaller than about λ. Hence, in any practical application, λ will always satisfy $\sigma_n \le \lambda \le \sigma_1$. An alternative method for regularization of (1) is the *truncated SVD* (TSVD) method, in which one discards the smallest singular values simply by truncating the sum in (5) at some $k < n$ [22]. Thus, the *TSVD solution* x_k is defined by

$$x_k \equiv \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i = U \Sigma_k^+ V^T b \ , \quad \Sigma_k^+ \equiv \mathrm{diag}(\sigma_1^{-1}, \cdots, \sigma_k^{-1}, 0, \cdots, 0) \ . \tag{8}$$

The integer k is called the *truncation parameter*, and it plays a role similar to the λ in Eq. (7). Notice that x_o is identical to x_λ with $\lambda = 0$ and x_k with $k = n$. We stress that the TSVD solution can be computed at least as efficiently as the regularized solution, without the large computational effort involved in a complete SVD computation, cf. the survey of methods in [11]. For more details, computational aspects, and examples of the application of these methods, see e.g. [2,3,5,6,8,22,23,24].

The use of Tikhonov regularization and TSVD is based on the following heuristic:

Heuristic 2.1. *The number of oscillations in the left and right singular vectors u_i and v_i tends to increase with increasing i.*

When this is true (which is the case e.g. if the matrix A is totally positive), it is obvious that the TSVD solution x_k as well as the regularized solution x_λ tend to be smoother than the least squares solution x_o. We are aware that in some applications, Heuristic 2.1 is not satisfied, or x_o is simply not the solution one is interested in (even without noise being present), because the sought solution f does not have a nice representation in terms of the right singular functions v_i. In these cases, one should replace the term $\|x\|_2$ in Eq. (6) by another appropriate regularization term such as e.g. $\|L\,x\|_2$, as pointed out in [23]. We are also aware that this corresponds to an expansion of the solution in terms of the generalized SVD of the matrix pair $(A\,,L)$ [12,23]. However, we feel that a fundamental understanding of the simpler case (6), as provided in this paper, is necessary before we can proceed to perform an analysis of the general case.

3. The convergence of the methods

Before starting our discussion, we shall make the assumption that k and λ are chosen such that the solutions x_λ and x_k are not too different — otherwise there will be no point in making a comparison between them. From the expression (7) for x_λ we see that this is the case when $\lambda \approx \sigma_k$, since then the damping of the terms in (7) sets in for singular values smaller than about σ_k. It can actually be proved [9, Theorem 5.2] that λ chosen somewhere in the range $(\sigma_k^3\,\sigma_{k+1})^{1/4} \le \lambda \le (\sigma_k\,\sigma_{k+1})^{1/2}$ brings x_k and x_λ as close as possible. In section 5 we return to the actual choice of k and λ.

The main goal of this section is to show how the behavior of the right-hand side b in (1) influences the convergence of the TSVD solution x_k and the regularized solution x_λ. For this purpose we set up, in terms of the SVD of A, the following 'model' of a right-hand side:

$$\beta_i = u_i^T b = \sigma_i^{\alpha}\ , \quad i = 1,\ldots,n\ , \quad \alpha \ge 0 \tag{9}$$

where the non-negative real constant α determines the decay of the β_i relative to the σ_i (when $\alpha > 1$, the β_i decay faster than the σ_i). According to (9), we can write $\boldsymbol{b} = U\,\Sigma^{\alpha} f$ with $f = [1, 1, \cdots, 1]^T$. Although Eq. (9) is a crude 'model' of the right-hand sides as they appear in practical applications, it is sufficiently realistic to clearly illustrate the importance of the decay of the β_i — rather than the particular shape of the singular values spectrum of A.

First, we investigate the convergence of the solutions $\boldsymbol{x}_k$ and $\boldsymbol{x}_\lambda$; i.e., we shall determine how well they approximate the exact least-squares solution $\boldsymbol{x}_o$ (5) to the unperturbed problem. Following [1], the differences $\boldsymbol{x}_o - \boldsymbol{x}_k$ and $\boldsymbol{x}_o - \boldsymbol{x}_\lambda$ are called the *TSVD error* and the *regularization error*, respectively. The closeness of $\boldsymbol{x}_k$ and $\boldsymbol{x}_\lambda$ to $\boldsymbol{x}_o$ is illustrated in the following theorem:

Theorem 3.1. *Let $\boldsymbol{x}_k$ and $\boldsymbol{x}_\lambda$ denote the solutions (8) and (7), and let $\boldsymbol{x}_o$ denote the exact least-squares solution (5). Further, let the right-hand side $\boldsymbol{b}$ satisfy Eq. (9). Then the norms of the TSVD error $\boldsymbol{x}_o - \boldsymbol{x}_k$ and the regularization error $\boldsymbol{x}_o - \boldsymbol{x}_\lambda$ satisfy:*

$$\frac{\|\boldsymbol{x}_o - \boldsymbol{x}_k\|_2}{\|\boldsymbol{x}_o\|_2} \le \begin{cases} \sqrt{n} & , \quad 0 \le \alpha < 1 \\ (\sigma_{k+1}/\sigma_1)^{\alpha-1}\sqrt{n} & , \quad 1 \le \alpha \end{cases} \tag{10a}$$

$$\frac{\|\boldsymbol{x}_o - \boldsymbol{x}_\lambda\|_2}{\|\boldsymbol{x}_o\|_2} \le \begin{cases} (\lambda/\sigma_n)^{\alpha-1}\sqrt{n} & , \quad 0 \le \alpha < 1 \\ (\lambda/\sigma_1)^{\alpha-1}\sqrt{n} & , \quad 1 \le \alpha\,. \end{cases} \tag{10b}$$

Proof. To simplify the notation, we introduce the quantities

$$\xi_o = V^T \boldsymbol{x}_o\,, \quad \xi_k = V^T \boldsymbol{x}_k\,, \quad \xi_\lambda = V^T \boldsymbol{x}_\lambda\,. \tag{11}$$

We shall first derive a lower bound on $\|\boldsymbol{x}_o\|_2$. According to (5):

$$\boldsymbol{x}_o = A^+ \boldsymbol{b} = V\,\Sigma^{-1} U^T U\,\Sigma^{\alpha} f = V\,\Sigma^{\alpha-1} f = V\,[\sigma_1^{\alpha-1}, \cdots, \sigma_n^{\alpha-1}]^T \quad \Leftrightarrow$$

$$\|\boldsymbol{x}_o\|_2 = \|\xi_o\|_2 \ge \max_i \{\sigma_i^{\alpha-1}\} = \begin{cases} \sigma_n^{\alpha-1} & , \quad 0 \le \alpha < 1 \\ \sigma_1^{\alpha-1} & , \quad 1 \le \alpha\,. \end{cases}$$

Concerning the norm of the TSVD error, we have

$$\|\boldsymbol{x}_o - \boldsymbol{x}_k\|_2 = \|\xi_o - \xi_k\|_2 \le \sqrt{n}\;\|\xi_o - \xi_k\|_\infty = \sqrt{n}\;\|(\Sigma^+ - \Sigma_k^+)\,\Sigma^{\alpha}\|_\infty$$

$$= \max_{k+1 \le i \le n} \{\sigma_i^{\alpha-1}\} \cdot \sqrt{n} = \begin{cases} \sigma_n^{\alpha-1}\sqrt{n} & , \quad 0 \le \alpha < 1 \\ \sigma_{k+1}^{\alpha-1}\sqrt{n} & , \quad 1 \le \alpha\,. \end{cases}$$

These relations immediately lead to Eq. (10a). To obtain a bound on the norm of the regularization error, we use the technique from [8, Chapter 2]. Defining the vector $\boldsymbol{w} \equiv V f$, it follows that the relation $b = A\,(A^TA)^\nu w \;\Leftrightarrow\; U\,\Sigma^\alpha f = U\,\Sigma^{2\nu+1} f$ is satisfied if $\nu = ½\,(\alpha-1)$. Thus, from [8, Theorem 2.2.2] we obtain

$$\|x_o - x_\lambda\|_2 \le \lambda^{2\nu}\,\|w\|_2 = \lambda^{\alpha-1}\sqrt{n}\ .$$

This relation, together with the bound for $\|x_o\|_2$, yields Eq. (10b). □

Theorem 3.1 shows that as long as α is larger than 1, and provided that σ_k and λ are small compared to $\sigma_1 = \|A\|_2$, both the TSVD solution x_k and the regularized solution x_λ are *guaranteed* to approximate x_o, and the larger α the better approximation. Usually, the truncation parameter k and the regularization parameter λ are determined by the errors in (1) in such a way that larger errors lead to a smaller k and a larger λ. Hence, Theorem 3.1 shows that if errors are present in (1), then x_k and x_λ can only yield satisfactory approximations to x_o if the coefficients β_i of the *unperturbed* right-hand side b decay to zero somewhat faster than the singular values σ_i. And the larger the errors, the faster the decay must be to ensure convergence.

Next, let us consider the similarity between the TSVD solution and the regularized solution by considering their difference $x_k - x_\lambda$ and also the difference between their residuals $(b - A\,x_k) - (b - A\,x_\lambda) = -A\,(x_k - x_\lambda)$. We assume that $\sigma_{k+1} \le \lambda \le \sigma_k$ and that the right-hand side b satisfies the 'model' (9). A convenient way to measure the difference between x_k and x_λ, as a function of α, is to define the following relative difference function:

$$\delta_k(\alpha) \equiv \min_{\sigma_{k+1}\le\lambda\le\sigma_k} \|V^T(x_k - x_\lambda)\|_\infty \cdot \frac{\|A\|_2}{\|U^Tb\|_\infty}\ . \tag{12a}$$

Similarly, we can measure the difference between the residuals by means of the function

$$\rho_k(\alpha) \equiv \min_{\sigma_{k+1}\le\lambda\le\sigma_k} \|U^TA\,(x_k - x_\lambda)\|_\infty / \|U^Tb\|_\infty\ . \tag{12b}$$

The distance function $\delta_k(\alpha)$ was briefly analyzed in [9, Theorem 6.1], and the analysis is extended in the following theorem.

Theorem 3.2. *Let x_k and x_λ denote the solutions (8) and (7), and let the right-hand side b satisfy Eq. (9). Then upper bounds for the functions $\delta_k(\alpha)$ and $\rho_k(\alpha)$ defined in Eqs. (12a) and (12b) are given by*

$$\delta_k(\alpha) \le \begin{cases} (\sigma_{k+1}/\sigma_1)^{\alpha-1} & , \quad \alpha < 3 \\ (\sigma_{k+1}/\sigma_1)^2 & , \quad \alpha \ge 3 \end{cases} \tag{13a}$$

$$\rho_k(\alpha) \le \begin{cases} (\sigma_{k+1}/\sigma_1)^{\alpha} & , \quad \alpha < 2 \\ (\sigma_{k+1}/\sigma_1)^2 & , \quad \alpha \ge 2 \end{cases} \tag{13b}$$

Proof. First, we notice that $V^T(x_k - x_\lambda) = \xi_k - \xi_\lambda$, where ξ_k and ξ_λ are defined in (11). From [9, Theorem 6.1] (in particular, the third column of Table 1) it follows that

$$\delta_k(\alpha) \le (\sigma_{k+1}/\sigma_1)^{\alpha-1} (\sigma_{k+1}/\sigma_k)^{\frac{1}{2}(3-\alpha)} \le (\sigma_{k+1}/\sigma_1)^{\alpha-1} \quad \text{for} \quad 0 \le \alpha \le 2$$

$$\delta_k(\alpha) \le (\sigma_{k+1}/\sigma_1)^2 \left[1 + (\sigma_{k+1}/\sigma_k)^2\right]^{-1} \le (\sigma_{k+1}/\sigma_1)^2 \quad \text{for} \quad \alpha \ge 3 \, .$$

For α in the interval $[2,3]$, a careful analysis of the norm

$$\|\xi_k - \xi_\lambda\|_\infty = \max \left\{ \frac{\sigma_1^{\alpha-1}\lambda^2}{\sigma_1^2+\lambda^2}, \ \cdots \ , \frac{\sigma_k^{\alpha-1}\lambda^2}{\sigma_k^2+\lambda^2}, \frac{\sigma_{k+1}^{\alpha+1}}{\sigma_{k+1}^2+\lambda^2}, \ \cdots \ , \frac{\sigma_n^{\alpha+1}}{\sigma_n^2+\lambda^2} \right\}$$

shows that its maximum value, for $\sigma_{k+1} \le \lambda \le \sigma_k$, is given by

$$\|\xi_k - \xi_\lambda\|_\infty = \begin{cases} F_\alpha \cdot \lambda^{\alpha-1} \, , & \sigma_{k+1} \le \lambda \le \left[\dfrac{3-\alpha}{\alpha-1}\right]^{\frac{1}{2}} \sigma_1 \\ \dfrac{\sigma_1^{\alpha-1}\lambda^2}{\sigma_1^2+\lambda^2} \, , & \left[\dfrac{3-\alpha}{\alpha-1}\right]^{\frac{1}{2}} \sigma_1 \le \lambda \le \sigma_k \end{cases} \tag{14}$$

where $F_\alpha \equiv \frac{1}{2}(\alpha-1)^{\frac{1}{2}(\alpha-1)}(3-\alpha)^{\frac{1}{2}(3-\alpha)}$. Since $\lambda^2/(\sigma_1^2+\lambda^2)$ increases with λ, it follows that the minimum of (14) for $\sigma_{k+1} \le \lambda \le \sigma_k$ is

$$\min_{\sigma_{k+1} \le \lambda \le \sigma_k} \|\xi_k - \xi_\lambda\|_\infty = \begin{cases} F_\alpha \cdot \sigma_{k+1}^{\alpha-1} \, , & \sigma_{k+1} \le \lambda \le \sigma_1 \left[\dfrac{3-\alpha}{\alpha-1}\right]^{\frac{1}{2}} \\ \dfrac{\sigma_1^{\alpha-1}\sigma_{k+1}^2}{\sigma_1^2+\sigma_{k+1}^2} \, , & \sigma_1 \left[\dfrac{3-\alpha}{\alpha-1}\right]^{\frac{1}{2}} \le \lambda \le \sigma_k \, . \end{cases}$$

Inserting this result into the expression (12a) for $\delta_k(\alpha)$, and noting that $\|A\|_2 / \|U^T b\|_\infty = \sigma_1^{1-\alpha}$ for all $\alpha \ge 0$, we obtain

$$\delta_k(\alpha) = \begin{cases} F_\alpha \cdot (\sigma_{k+1}/\sigma_1)^{\alpha-1} \, , & \sigma_{k+1} \le \lambda \le \sigma_1 \left[\dfrac{3-\alpha}{\alpha-1}\right]^{\frac{1}{2}} \\ (\sigma_{k+1}/\sigma_1)^2 \left[1 + (\sigma_{k+1}/\sigma_1)^2\right]^{-1} \, , & \sigma_1 \left[\dfrac{3-\alpha}{\alpha-1}\right]^{\frac{1}{2}} \le \lambda \le \sigma_k \, . \end{cases}$$

Here, $F_\alpha \le 1$ and $[1 + (\sigma_k/\sigma_1)^2]^{-1} \le 1$, and since $(\sigma_{k+1}/\sigma_1)^{\alpha-1} > (\sigma_{k+1}/\sigma_1)^2$ for $2 \le \alpha \le 3$, an upper bound for $\delta_k(\alpha)$ is $(\sigma_{k+1}/\sigma_1)^{\alpha-1}$. This establishes (13a). To prove (13b), we notice that

$\|U^T b\|_\infty^{-1} = \sigma_1^{-\alpha}$ and that $U^T A\,(x_k - x_\lambda) = \Sigma\,(\xi_k - \xi_\lambda)$. Hence, $\rho_k(\alpha) = \delta_k(\alpha-1)$, and (13b) therefore follows from (13a) with α replaced by $\alpha+1$. □

Theorem 3.2 shows that when k is fixed by the errors present in (1) and σ_{k+1} is small compared to σ_1, then there exists a $\lambda \in [\sigma_{k+1}, \sigma_k]$ such that x_k and x_λ, as well as the corresponding residulas, are *guaranteed* to be close whenever α is larger than 1. And the larger α the closer the solutions and the residuals. This means that whenever the coefficients β_i of the *unperturbed* right-hand side b decay to zero somewhat faster than the singular values σ_i, then TSVD and Tikhonov regularization will produce approximately the same solutions and residuals, and according to Theorem 3.1 both of the solutions will approximate the unperturbed least-squares solution x_o.

We conclude this section by giving a more rigorous definition of the requirement on b. Of course, the decay of the β_i-coefficients need not be monotonic, as long as the β_i *in average* decay to zero faster than the σ_i. We can formulate this requirement as follows:

Definition 3.3. The discrete Picard condition (DPC). *Let b denote an unperturbed right-hand side in (1). Then b satisfies the DPC if*

$$\frac{|u_{i+1}^T b|}{\sigma_{i+1}} \le C\,\frac{|u_i^T b|}{\sigma_i}\,, \quad i = 1, \cdots, n-1 \tag{15}$$

where C is a constant of order unity. If some $u_i^T b / \sigma_i$ is numerically zero, it should be replaced by the first previous nonzero $u_{i-j}^T b / \sigma_{i-j}$, $j = 1, 2, \cdots$.

We remark that if the discrete problem (1) is obtained from the integral equation (2) by means of an expansion method with orthonormal basis functions, and if the integral equation satisfies the Picard condition (4), then the DPC is also satisfied due to the relationship between the SVE and the SVD [10].

It should be stressed that while the *convergence* of x_k and x_λ depends on the DPC, the *smoothness* of these solutions depends on Heuristic 2.1. Thus, both vectors x_k and x_λ may be smooth even if the DPC is not satisfied, but in that case they will not approximate the vector x_o. Although such solutions may still be acceptable in certain cases, it would be more correct to use the general formulation of regularization as mentioned in the last paragraph of Section 2.

4. Perturbation bounds and condition numbers

In many discrete ill-posed problems, the errors are restricted to the right-hand side only. For example, this is the case if (1) is derived from an integral equation (2) whose kernel K is given exactly, e.g. from some mathematical model of a physical problem, while the right-hand side consists of measured quantities contaminated by errors. These errors transform directly into a perturbation of the right-hand side $\boldsymbol{b}$ in (1). Examples of such problems are inverse problems in observational astronomy [3], the inverse problem of electrocardiography [4], deconvolution problems such as inverse Radon and inverse Laplace transforms [15,24], and inverse problems in computational physics (see [16] for an overview and [6] for a specific example). It is therefore appropriate to derive bounds on the perturbations of $\boldsymbol{x}_k$ and $\boldsymbol{x}_\lambda$ solely due to a perturbation $\boldsymbol{e}$ of $\boldsymbol{b}$.

Theorem 4.1. *Let $\boldsymbol{x}_k$ and $\boldsymbol{x}_\lambda$ denote the solutions (8) and (7), and let $\bar{\boldsymbol{x}}_k$ and $\bar{\boldsymbol{x}}_\lambda$ denote the solutions when the right-hand side $\boldsymbol{b}$ of (1) is perturbed by $\boldsymbol{e}$. Assuming that $\sigma_n \le \lambda \le \sigma_1$, the relative perturbations are bounded as:*

$$\frac{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}_k\|_2}{\|\boldsymbol{x}_k\|_2} \le \frac{\sigma_1}{\sigma_k} \frac{\|\boldsymbol{e}\|_2}{\|\boldsymbol{b}_k\|_2} \tag{16a}$$

$$\frac{\|\boldsymbol{x}_\lambda - \bar{\boldsymbol{x}}_\lambda\|_2}{\|\boldsymbol{x}_\lambda\|_2} \le \frac{\sigma_1}{2\lambda} \frac{\|\boldsymbol{e}\|_2}{\|\boldsymbol{b}_\lambda\|_2} \tag{16b}$$

where $\boldsymbol{b}_k = A\,\boldsymbol{x}_k$ and $\boldsymbol{b}_\lambda = A\,\boldsymbol{x}_\lambda$.

Proof. It is elementary to derive (16a) from the inequalities $\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}_k\|_2 = \|V\,\Sigma_k^+ U^T \boldsymbol{e}\|_2 \le \|\Sigma_k^+\|_2 \|\boldsymbol{e}\|_2 = \sigma_k^{-1} \|\boldsymbol{e}\|_2$ and $\|\boldsymbol{b}_k\|_2 = \|A\,\boldsymbol{x}_k\|_2 \le \|A\|_2 \|\boldsymbol{x}_k\|_2 = \sigma_1 \|\boldsymbol{x}_k\|_2$. Further, it follows that $\|\boldsymbol{x}_\lambda - \bar{\boldsymbol{x}}_\lambda\|_2 = \|V\,\Sigma_\lambda^+ U^T \boldsymbol{e}\|_2 \le \|\Sigma_\lambda^+\|_2 \|\boldsymbol{e}\|_2$, and assuming that $\sigma_n \le \lambda \le \sigma_1$ one obtains

$$\|\Sigma_\lambda^+\|_2 = \max_{1 \le i \le n} \left\{ \frac{\sigma_i}{\sigma_i^2 + \lambda^2} \right\} \le \max_{\sigma_n \le \sigma \le \sigma_1} \frac{\sigma}{\sigma^2 + \lambda^2} = 1/(2\lambda)\,.$$

These two relations, together with $\|\boldsymbol{b}_\lambda\|_2 \le \sigma_1 \|\boldsymbol{x}_\lambda\|_2$, lead directly to (16b). □

Remark. The bound in Eq. (16a) can also be derived from Lemma 2.3.2 in [8]. Applying the same lemma to Tikhonov regularization, we obtain a factor σ_1/λ instead of the factor $\sigma_1/(2\lambda)$ in Eq. (16b). If the error norms are bounded relative to $\|\boldsymbol{x}_o\|_2$ instead of $\|\boldsymbol{x}_k\|_2$ and $\|\boldsymbol{x}_\lambda\|_2$, the right-hand sides simply change to $\sigma_1/\sigma_n\ \|\boldsymbol{e}\|_2/\|\boldsymbol{b}_o\|_2$ and $\sigma_1/(2\lambda)\ \|\boldsymbol{e}\|_2/\|\boldsymbol{b}_o\|_2$, where $\boldsymbol{b}_o = A\,\boldsymbol{x}_o$.

Theorem 4.1 confirms what has been observed experimentally and used in a number of applications, namely that it is possible to choose k and λ such that the approximate perturbed solutions $\bar{\boldsymbol{x}}_k$ and

$\bar{x}_\lambda$ are fairly insensitive to the perturbations in b. The theorem also shows that when $\lambda \approx \sigma_k$, as we assumed in the previous section, then both methods are approximately equally sensitive to the perturbations. Notice that there is always a tradeoff between Theorem 4.1 and Theorem 3.1 in the choice of k and λ: when k is small and λ is large then the perturbation bounds are small while the TSVD and regularization errors may be large, and vice versa.

The results in Theorem 4.1 can also be used to derive expressions for the *condition numbers* associated with TSVD and Tikhonov regularization. Here, we shall use the following definitions:

Definition 4.2. *The condition numbers κ_k and κ_λ, associated with TSVD and Tikhonov regularization, respectively, are defined as*

$$\kappa_k \equiv \lim_{\|e\|_2 \to 0} \sup \frac{\|x_k - \bar{x}_k\|_2}{\|x_k\|_2} \,, \quad \kappa_\lambda \equiv \lim_{\|e\|_2 \to 0} \sup \frac{\|x_\lambda - \bar{x}_\lambda\|_2}{\|x_\lambda\|_2} \tag{17}$$

where x_k and x_λ are the unperturbed TSVD and regularized solutions, and $\bar{x}_k$ and $\bar{x}_\lambda$ are the solutions when the right-hand side is perturbed by e.

This definition, together with Theorem 4.1, immediately leads to:

Corollary 4.3. *The condition numbers (17) associated with TSVD and Tikhonov regularization are*

$$\kappa_k = \sigma_1 / \sigma_k \,, \quad \kappa_\lambda = \sigma_1 / (2\lambda) \,. \tag{18}$$

Remark. Schock [19] recently considered another condition number $\bar{\kappa}_\lambda = \lambda / \sigma_n$ associated with Tikhonov regularization, based on the usual condition number of the matrix $A_\lambda^I = (A^T A + \lambda^2 I)^{-1} A^T$, which is the unique matrix that produces the regularized solution $x_\lambda = A_\lambda^I b$. The result λ / σ_n is, however, not quite right. Instead it should be:

$$\bar{\kappa}_\lambda = \begin{cases} \sigma_1 / \lambda & , \quad \lambda \le \sqrt{\sigma_1 \sigma_n} \\ \lambda / \sigma_n & , \quad \lambda > \sqrt{\sigma_1 \sigma_n} \,. \end{cases}$$

which is easily derived from the proof for [19, Theorem 2]. We feel that our condition number κ_λ (18) is more correct than $\bar{\kappa}_\lambda$ since the matrix A_λ^I should not be used to compute x_λ numerically.

Next, we shall give the general perturbation bounds for x_k and x_λ when both the matrix A and the right-hand side b are perturbed. The perturbation of A may e.g. arise from the approximations used to derive A from the kernel K in the numerical treatment of the integral equation (2).

Theorem 4.4. *Let E and e denote the perturbations of A and b, respectively, and assume that $\|E\|_2 < \sigma_k - \sigma_{k+1}$ and $\|E\|_2 < \lambda$. Then the perturbations of x_k and x_λ are bounded by*

$$\frac{\|x_k - \bar{x}_k\|_2}{\|x_k\|_2} \le \frac{\|A\|_2}{\sigma_k - \|E\|_2} \left[\left[2 + \frac{\sigma_{k+1}}{\omega_k}\right] \frac{\|E\|_2}{\|A\|_2} + \frac{\|e\|_2}{\|b_k\|_2} + \frac{\|E\|_2}{\omega_k} \frac{\|r_k\|_2}{\|b_k\|_2} \right] \qquad (19a)$$

$$\frac{\|x_\lambda - \bar{x}_\lambda\|_2}{\|x_\lambda\|_2} < \frac{\lambda + \|A\|_2}{\lambda - \|E\|_2} \left[2 \frac{\|E\|_2}{\|A\|_2} + \frac{\|e\|_2}{\|b_\lambda\|_2} + \frac{\|E\|_2}{\lambda} \frac{\|r_\lambda\|_2}{\|b_\lambda\|_2} \right] \qquad (19b)$$

where $r_k = b - A\, x_k$, $r_\lambda = b - A\, x_\lambda$, *and* $\omega_k = \sigma_k - \sigma_{k+1} - \|E\|_2$.

Proof. Eq. (19a) follow immediately from [9, Theorem 3.4] together with the relation

$$\frac{\|E\|_2/\sigma_k}{1 - \|E\|_2/\sigma_k - \sigma_{k+1}/\sigma_k} = \frac{\|A\|_2}{\sigma_k - \|E\|_2} \frac{\sigma_k - \|E\|_2}{\omega_k} \frac{\|E\|_2}{\|A\|_2} = \frac{\|A\|_2}{\sigma_k - \|E\|_2} \left[1 + \frac{\sigma_{k+1}}{\omega_k}\right] \frac{\|E\|_2}{\|A\|_2} .$$

To prove Eq. (19b), we remind that x_λ is the unique least squares solution to the problem $\min \|C\,x - d\|_2$, where $C = \begin{bmatrix} A \\ \lambda I \end{bmatrix}$ and $d = \begin{bmatrix} b \\ 0 \end{bmatrix}$. The matrix C has full rank for all $\lambda > 0$, and we can apply the standard perturbation bound for least squares solutions [2, Thm. 5.5] to get:

$$\|x_k - \bar{x}_\lambda\|_2 \le \frac{\|C\|_2 \|C^+\|_2}{1 - \|E\|_2 \|C^+\|_2} \left[\frac{\|E\|_2}{\|C\|_2} \|x_\lambda\|_2 + \frac{\|e\|_2}{\|C\|_2} + \|E\|_2 \|C^+\|_2 \frac{\|d - C\,x_\lambda\|_2}{\|C\|_2} \right] .$$

From the definitions of C and d we get $\|A\|_2 < \|C\|_2 \le \|A\|_2 + \lambda$, $\|C^+\|_2 = \sigma_n(C)^{-1} < \lambda^{-1}$, $\|C\|_2 \|x_\lambda\|_2 \ge \|C\,x_\lambda\|_2 > \|A\,x_\lambda\|_2 = \|b_\lambda\|_2$ and $\|d - C\,x_\lambda\|_2 = \left\| \begin{bmatrix} r_\lambda \\ \lambda x_\lambda \end{bmatrix} \right\|_2 \le \|r_\lambda\|_2 + \lambda\, \|x_\lambda\|_2$ such that

$$\frac{\|C\|_2 \|C^+\|_2}{1 - \|E\|_2 \|C^+\|_2} < \frac{\lambda + \|A\|_2}{\lambda - \|E\|_2} , \qquad \frac{\|E\|_2}{\|C\|_2} < \frac{\|E\|_2}{\|A\|_2} , \qquad \frac{\|e\|_2}{\|C\|_2 \|x_\lambda\|_2} < \frac{\|e\|_2}{\|b_\lambda\|_2}$$

$$\|E\|_2 \|C^+\|_2 \frac{\|d - C\,x_\lambda\|_2}{\|C\|_2 \|x_\lambda\|_2} < \frac{\|E\|_2}{\lambda} \frac{\|r_\lambda\|_2}{\|b_\lambda\|_2} + \frac{\|E\|_2}{\|A\|_2} .$$

These relations then yield Eq. (19b). □

Although the bound in Eq. (19b) is not tight, it does illustrate our major point, namely that the general perturbation bounds for x_k and x_λ are very similar whenever $\lambda \approx x_k$ (which we have already assumed), provided that k and λ are chosen such that $\|E\|_2 < \sigma_k - \sigma_{k+1}$ and $\|E\|_2 < \lambda$. We see from Eq. (19a) that truncation of the sum in (8) at a nearly multiple singular value σ_k should be avoided; but apart from this, the results in Theorem 4.1 do not impose any particular requirement on the singular value spectrum. I.e., one can actually truncate the expression (8) for x_k at any value of k, as long as

σ_k is not nearly multiple.

The main conclusion to be drawn from Theorems 3.1, 3.2, 4.1 and 4.4 is therefore that the success of TSVD (as well as Tikhonov regularization) primarily depends on satisfaction of the DPC (15) and in fact has little to do with the existence of a gap in the singular value spectrum of A. If, for some k, there is a large gap between σ_k and σ_{k+1} (i.e. A has well-determined numerical rank), then this k is usually identical to the numerical rank of A and it is therefore often convenient to truncate the expression for x_k at this k [9]. In this case, it is natural to require the DPC satisfied for the first k coefficients β_i only, and to consider the remaining β_i-coefficients associated with the residual. If, on the other hand, A has ill-determined numerical rank then there is no point in trying to find A's numerical rank. Instead, one should choose k in order to suppress, as much as possible, the influence of the perturbations while, at the same time, keeping the TSVD error as small as possible. In the next section, we shall discuss this in more details.

5. Characterization of the solutions

In this section we are mainly interested in the TSVD solution and the regularized solution when they are influenced by perturbations of the right-hand side. In order to understand the influence of such perturbations and to be able to select a proper truncation parameter k and regularization parameter λ, we therefore seek to characterize the behavior of the perturbed solutions $\tilde{x}_k$ and $\tilde{x}_\lambda$ as functions of k and λ. A convenient way to characterize any solution to the least squares problem (1) is to plot its norm versus the norm of the corresponding residual, as suggested in [13, Chapter 26]. Since we are only interested in that component of the residual which lies in the column space of A, we define the residuals r_k and r_λ corresponding to x_k and x_λ by

$$r_k \equiv b_o - A\,x_k = \sum_{i=k+1}^{n} u_i^T b\; u_i \;, \quad r_\lambda \equiv b_o - A\,x_\lambda = \sum_{i=1}^{n} \frac{\lambda^2}{\sigma_i^2 + \lambda^2}\, u_i^T b\; u_i \;. \tag{20}$$

It can be proved that for regularization $\|x_\lambda\|_2$ is a decreasing function of $\|r_\lambda\|_2$, while for TSVD $\|x_k\|_2$ is a decreasing function of $\|r_k\|_2$ on a finite set [10, Theorem 5.3], and that the points $(\|r_k\|_2, \|x_k\|_2)$, $k=1,\ldots,n-1$ always lie above the curve $(\|r_\lambda\|_2, \|x_\lambda\|_2)$ [13, Theorem (25.49)]. The distance between these points and the curve was already analyzed in Theorem 3.2. Here, we shall give a more detailed (although not strictly rigorous) description of the curve and the set of points.

First, let us consider the behavior of the curve $(\|r_\lambda\|_2, \|x_\lambda\|_2)$, with x_λ and r_λ given by Eqs. (7) and (20). Obviously, $x_\lambda \to 0$ and $r_\lambda \to b_o$ as $\lambda \to \infty$, while $x_\lambda \to x_o$ and $r_\lambda \to 0$ as $\lambda \to 0$. For

$\lambda \ll \sigma_n$, we have

$$\sigma_i^2 + \lambda^2 \approx \sigma_i^2 \quad \Rightarrow \quad \|r_\lambda\|_2 \approx \lambda^2 \left[\sum_{i=1}^{n} \left[u_i^T b / \sigma_i^2 \right]^2 \right]^{1/2} \le (\lambda / \sigma_n)^2 \, \|b_o\|_2$$

and $\sigma_i^2/(\sigma_i^2+\lambda^2) \approx 1 \Rightarrow x_\lambda \approx x_o$; i.e., for small λ the curve $(\|r_\lambda\|_2, \|x_\lambda\|_2)$ is approximately a horizontal line at $\|x_\lambda\|_2 \approx \|x_o\|_2$. When λ increases, we see that $\|x_\lambda\|_2$ starts to decrease while $\|r_\lambda\|_2$ still grows towards $\|b_o\|_2$, and thus the curve must bend down towards the abscissa axis. The value of λ for which $\|x_\lambda\|_2$ markedly starts to bend down depends on the σ_i as well as the $u_i^T b$. If the DPC is satisfied, such that the sum in the expression (7) for x_λ is dominated by its first terms, then obviously λ must be comparable with the largest singular values σ_i to significantly influence x_λ. For such values of λ, $\|r_\lambda\|_2$ will be somewhat smaller than $\|b_o\|_2$ because the coefficients to the first terms in the expression (20) for r_λ will be less than one. — If the DPC is not satisfied, we can assume that most of the terms in the expression for x_λ actually contribute to this vector, and the influence of λ can be felt for much smaller values of λ than before. Thus, the curve also starts to bend down for smaller values of $\|r_\lambda\|_2$ than before.

Next, we consider the points $(\|r_k\|_2, \|x_k\|_2)$. If the DPC is satisfied, Theorem 3.2 guarantees that for large k there always exists a $\lambda \in [\sigma_{k+1}, \sigma_k]$ such that the points are close to the curve $(\|r_\lambda\|_2, \|x_\lambda\|_2)$, while for small k we cannot guarantee this. However, we can see from the expressions (8) and (20) for x_k and r_k that their norms must behave quantitatively like $\|x_\lambda\|_2$ and $\|r_\lambda\|_2$. I.e., if the DPC is satisfied, then for large k the points will approximately be on the same horizontal line as the curve, and as k gets smaller the points will eventually start to bend down, always lying strictly above the curve. If the DPC is not satisfied, the points will deviate from the curve even for large k.

If we make more assumptions about the case when the right-hand side does not satisfy the DPC, we can also say more about the curve and the points. This is particularly relevant for the perturbation e of b. An interesting case (see below) is when all the coefficients $|u_i^T e|$ of the perturbation are approximately of the same size,

$$|u_i^T e| \approx \varepsilon_o \, , \quad i = 1, \cdots, m \, .$$

The corresponding 'solution' $x_\lambda^{(e)}$ and 'residual' $r_\lambda^{(e)}$ are given by

$$x_\lambda^{(e)} \approx \varepsilon_o \sum_{i=1}^{n} \frac{\sigma_i}{\sigma_i^2 + \lambda^2} \, v_i \, , \quad r_\lambda^{(e)} \approx \varepsilon_o \sum_{i=1}^{n} \frac{\lambda^2}{\sigma_i^2 + \lambda^2} \, u_i \, . \tag{21}$$

The vector $x_\lambda^{(e)}$ is called the 'noise amplification error' in [1]. In particular the norm of $x_o^{(e)} = A^+e$ satisfies $\varepsilon_o / \sigma_n \le \|x_o^{(e)}\|_2 \le \sqrt{n}\ \varepsilon_o / \sigma_n$. We see from (21) that for λ in the range $\sigma_n \le \lambda < \infty$, $\|r_\lambda^{(e)}\|_2$ varies in the quite small range from approximately ε_o to $\sqrt{n}\ \varepsilon_o$. Concerning $x_\lambda^{(e)}$, the sum in (21) is dominated by just a few, say p, of the terms, namely those for which $\sigma_i \approx \lambda$ and $\sigma_i / (\sigma_i^2 + \lambda^2) \approx 1/(2\lambda)$ such that $\|x_\lambda^{(e)}\|_2 \approx p \cdot \varepsilon_o / (2\lambda) \approx \varepsilon_o / \lambda$. Thus, as $\lambda \to \infty$, the curve $(\|r_\lambda^{(e)}\|_2, \|x_\lambda^{(e)}\|_2)$ soon becomes almost a vertical line at $\|r_\lambda^{(e)}\|_2 \approx \sqrt{n}\ \varepsilon_o$. Exactly the same conclusions hold for the points $(\|r_k^{(e)}\|_2, \|x_k^{(e)}\|_2)$ obtained when applying TSVD to e.

Typical examples of both curves $(\|r_\lambda\|_2, \|x_\lambda\|_2)$, with the DPC satisfied, and $(\|r_\lambda^{(e)}\|_2, \|x_\lambda^{(e)}\|_2)$, with the DPC not satisfied, are shown in Fig. 1 for the situation $\|e\|_2 < \|b_o\|_2$. Notice that the latter curve starts to bend down towards the abscissa axis before the first curve does (e.g. for smaller λ). Also notice that the level $\|\bar{x}_o\|_2 \approx \|x_o^{(e)}\|_2 \approx \sqrt{n}\ \varepsilon_o / \sigma_n$ lies over the level $\|x_o\|_2$.

We are now ready to describe the behavior of the solutions $\bar{x}_\lambda = x_\lambda + x_\lambda^{(e)}$ and $\bar{x}_k = x_k + x_k^{(e)}$ under the influence of errors e in the right-hand side. We make the following assumptions in order to be able to carry out a meaningful analysis:

Assumption 5.1. *Let e be a perturbation of the right-hand side b in (1). We assume that*

1. *the unperturbed right-hand side b satisfies the DPC (15),*
2. *the perturbation e is a random vector of zero mean and covariance matrix $\sigma^2 I$,*
3. *the expected value of all coefficients $u_i^T e$ is a constant ε_o independently of i, and*
4. *the norm of e satisfies $\|e\|_2 < \|b_o\|_2$.*

The first assumption is necessary for the convergence of the methods. The second assumption is very common in least squares problems. If it is not satisfied, one should either scale the equations (if the covariance matrix is diagonal) or use the general Gauss-Markoff linear model for a general covariance matrix, cf. e.g. [2, Section 14] and [26]. The third assumption means that the errors are equally likely to contribute to all the terms in the sum (5) for x_o, which is also very common whenever Heuristic 2.1 is satisfied. Assumption 4 is simply a requirement that the signal-to-noise ratio in the given right-hand side is not too large to let one retrieve a satisfactory approximate solution.

With these assumptions it then follows from the above analysis that the curve $(\|\bar{r}_\lambda\|_2, \|\bar{x}_\lambda\|_2)$ and the set of points $(\|\bar{r}_k\|_2, \|\bar{x}_k\|_2)$ will appear as shown in Fig. 2. For $\lambda \ll \sigma_n$, the curve is almost horizontal at the level $\|\bar{x}_\lambda\|_2 \approx \|x_o^{(e)}\|_2 \approx \sqrt{n}\ \varepsilon_o / \sigma_n$, since the regularized solution $\bar{x}_\lambda$ is dominated by the term $x_\lambda^{(e)}$. Increasing λ, the curve soon starts to bend down due to the influence of λ in the term $x_\lambda^{(e)}$ which still dominates $\bar{x}_\lambda$. This part of the curve therefore resembles the dashed line in Fig. 1. Increasing λ further, the other term x_λ will start to dominate $\bar{x}_\lambda$ at some point such that the curve now stays at a new level at $\|\bar{x}_\lambda\|_2 \approx \|x_o\|_2$, until it eventually starts to bend down again for λ comparable with the largest singular values. This part of the curve is therefore similar to the solid line in Fig. 1. As long as Assumption 5.1 is satisfied, there will always be a more or less distinct 'corner' somewhere in the middle of the curve, where the dominating term in $\bar{x}_\lambda$ switches from x_λ to $x_\lambda^{(e)}$. The set of points corresponding to the TSVD solution $\bar{x}_k$ behaves in exactly the same way, also exhibiting a 'corner' when the dominating term in $\bar{x}_k$ switches from x_k to $x_k^{(e)}$. At this corner, and to its right, the component x_k dominates and Theorem 3.2 ensures that the points will be close to the solid curve (except for the smallest k). To the left, $x_k^{(e)}$ dominates, and as k increases the points will deviate from the solid curve. We can summarize the main result as follows:

Characterization 5.2. *If Assumption 5.1 is satisfied, then the curve* $(\|\bar{r}_\lambda\|_2, \|\bar{x}_\lambda\|_2)$ *as well as the set of points* $(\|\bar{r}_k\|_2, \|\bar{x}_k\|_2)$ *exhibit a 'corner' behavior as functions of their parameters* λ *and* k. *Both 'corners' occur approximately at* $(\sqrt{n}\ \varepsilon_o, \|x_o\|_2)$. *The larger the difference between the decay rates of* $|u_i^T b|$ *and* $|u_i^T e|$, *the more distinct the 'corners' will appear.*

The plots in Fig. 2 provide a natural choice of the regularization parameter λ and the truncation parameter k that must be selected. It is obvious that the optimal values of λ and k are those that yield solutions near the 'corners' of the curve and the point set, respectively, since these solutions approximate x_o as close as possible without being dominated by the contributions from the perturbation e. In the case of the TSVD solution $\bar{x}_k$, we can also say that k should be chosen as large as possible, but with the constraint that the k coefficients $u_i^T(b+e)$, $i = 1, \cdots, k$ in the truncated sum for $\bar{x}_k$ satisfy the DPC. The optimal solutions, produced by these optimal values of λ and k, satisfy $\|\bar{x}_\lambda\|_2 \approx \|\bar{x}_k\|_2 \approx \|x_o\|_2$ and $\|\bar{r}_\lambda\|_2 \approx \|\bar{r}_k\|_2 \approx \|e\|_2$; i.e., they are *reasonable solutions* in the terminology of Varah [24]. We stress that whenever Assumption 5.1 is satisfied, and λ and k are chosen as described above, then both TSVD and Tikhonov regularization are guaranteed to produce very similar reasonable solutions that *converge* to x_o as $e \to 0$.

We shall not go into a detailed description of numerical methods for determining λ and k according to the above criteria. Suffice it to say that the key idea in most of these methods is actually to locate the 'corners' of the curve and the point set as illustrated in Fig. 2. This is clearly the case in methods based on the *discrepancy principle* [8, Section 3.3] where one increases k or decreases λ until the residual norm (or some function of this norm) is of the same size as the norm of the errors $\|e\|_2 \approx \sqrt{n}\ \varepsilon_o$. *Generalized cross-validation* [7] is a promising alternative method which, in the case of TSVD, simply amounts to choosing k so as to minimize the function $\|\bar{r}_k\|_2^2/(m-n)^2$. This k is identical to the k for which $\|\bar{r}_k\|_2^2/(m-n)$, as a function of k, starts to level off and become an estimate of the variance $\|e\|_2^2/m$ of the noise, and the corresponding point $(\|\bar{r}_k\|_2, \|\bar{x}_k\|_2)$ is therefore near the 'corner' of the point set. The same holds for Tikhonov regularization, cf. the discussion in [11].

6. Numerical examples

This section includes two examples of the numerical solution of first kind Fredholm integral equations (2), illustrating the discussion in the previous sections. In both examples we choose $m = n$, and we discretize by means of the method of moments with simple orthonormal basis functions ϕ_i chosen to give a piecewise constant approximation to f:

$$\phi_i(x) = \begin{cases} h^{-1/2} & , \quad a+(i-1)h \le x \le a+i\,h \\ 0 & , \quad \text{otherwise} \end{cases} \qquad i = 1, \cdots, n$$

where $h = (b-a)/n$ and $[a, b]$ is the integration interval. The elements of A and b in the least squares problem (1) are then given by the integrals

$$a_{ij} = h^{-1} \int_{a+(i-1)h}^{a+ih} \int_{a+(j-1)h}^{a+jh} K(s,x)\,dx\,ds \ , \qquad b_i = h^{-1/2} \int_{a+(i-1)h}^{a+ih} g(s)\,ds \ . \tag{22}$$

For more details on this method see e.g. [10] where it is shown that the singular values σ_i of A, when computed by Eq. (22), are $O(n^{-2})$-approximations to the μ_i in the SVE of K.

The first example is a classical example by Phillips [17]. The integral equation is given by the following K and g:

$$K(s,x) = \begin{cases} 1 + \cos[\pi(s-x)/3] & , \quad |s-x| \le 3 \\ 0 & , \quad |s-x| > 3 \end{cases}$$

$$g(s) = (6 - |s|)\left[1 + \tfrac{1}{2}\cos(\pi s/3)\right] + \frac{9}{2\pi}\sin(\pi |s| / 3)$$

with $[a, b] = [-6, 6]$. The solution is $f(x) = 1 + \cos(\pi x/3)$. The explicit formulas for A and b, evaluated by means of (22), are given in [10]. We used $n = 64$ and perturbed the right-hand side b by random numbers from a normal distribution with zero mean and standard deviation 10^{-4} such that $\|e\|_2 \simeq 8\cdot 10^{-4}$. The condition number of A is $\sigma_1/\sigma_n = 2.8\cdot 10^5$, and we know that Assumption 5.1 is satisfied. For $i \le 13$ the computed quantities σ_i, $|u_i^T b|$ and $|u_i^T b|/\sigma_i$ (not shown here) all decay as expected. For $i \ge 13$ the singular values σ_i continue to decrease while the coefficients $|u_i^T b|$ settle at the error level about 10^{-4}. Hence, the $|u_i^T b|/\sigma_i$ increase almost monotonically with i for $i \ge 13$. Fig.

3 shows the curve $(\|\bar{r}_\lambda\|_2, \|\bar{x}_\lambda\|_2)$ and some of the points $(\|\bar{r}_k\|_2, \|\bar{x}_k\|_2)$; notice the likeness with Fig. 2. It is clear, both from these plots and from plots of $|u_i^T b| / \sigma_i$, that one should truncate the TSVD solution x_k at about $k = 12$. The approximate solution, computed from $\bar{x}_k$, agrees with the true f within a maximum deviation of about 10^{-2} which is largely due to the TSVD error $x_o - x_k$.

The second example is a real problem from observational astronomy, where the right-hand side g is the probability density function of *observed* stellar parallaxes, while f is the true probability density function of these parallaxes. Assuming that the measurement errors are normally distributed, it is easy to show that f is related to g by a first kind Fredholm integral equation (2) with

$$K(s,x) = \frac{1}{\rho\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}\frac{(s-x)^2}{\rho^2}\right] .$$

The factor ρ reflects the accuracy of the measurements. As a case study we used a standard set of observations from [20, Table 4, p. 30] defining g in the form of a piecewise constant function and with $\rho = 0.014234$. We used the interval $[a, b] = [-0.03, 0.10]$, and the elements of A (22) were calculated using the 2-dim. 9-point Simpson quadrature rule. The computed values of σ_i and $|u_i^T b| / \sigma_i$ are shown in Fig. 4. Notice that the $|u_i^T b| / \sigma_i$ initially decay slightly with i, and that they soon start to increase dramatically. It is evident from this figure that the TSVD should be truncated at k equal to 6 or 7. Fig. 5 shows plots of $(\|\bar{r}_\lambda\|_2, \|\bar{x}_\lambda\|_2)$ and $(\|\bar{r}_k\|_2, \|\bar{x}_k\|_2)$, and the similarity with the idealized plot in Fig. 2 indicates that Assumption 5.1 is actually satisfied. We see from Fig. 5 that the error level is about $\|e\|_2 \approx 10^{-3}$, that there is a distinct 'corner' on the curve, and that the TSVD solutions x_k with $k = 5, 6, 7, 8$ are close to this 'corner'. The computed solution corresponding to $k = 6$ turned out to give the best results.

Acknowledgement

I would like to thank Prof. Leslie Foster for discussions which led my investigations in the direction presented here.

References

[1] C.M. Aulick & T.M. Gallie, *Isolating error effects in solving ill-posed problems*, SIAM J. Alg. Disc. Meth. 4 (1983), 371-376.

[2] Å. Björck, *Least Squares Methods;* in P.G. Ciarlet & J.L. Lions (Eds.), *Handbook of Numerical Analysis, Vol. III: Finite Difference Methods — Solution of Equations in* $\mathbb{R}^n$, Elsevier (1989), to appear.

[3] I.J.D. Craig & J.C. Brown, *Inverse Problems in Astronomy*, Adam Hilger (1986).

[4] J.J.M. Cuppen, *Calculating the isochrones of ventricular depolarization*, SIAM J. Sci. Stat. Comput. 5 (1984), 105-120.

[5] F.R. de Hoog, *Review of Fredholm equations of the first kind;* in R.S. Anderssen, F.R. de Hoog & M.A. Lukas (Eds.), *The Application and Numerical Solution of Integral Equations*, Sijthoff & Noordhoff (1980), pp. 119-134.

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

[7] G.H. Golub, M.T. Heath & G. Wahba, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics 21 (1979), 215-223.

[8] C.W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Integral Equations of the First Kind*, Pitman (1984).

[9] P.C. Hansen, *The truncated SVD as a method for regularization*, BIT 27 (1987), 534-553.

[10] P.C. Hansen, *Computation of the singular value expansion*, Computing 40 (1988), 185-199.

[11] P.C. Hansen, *Solution of ill-posed problems by means of truncated SVD;* in (R.P. Agarwal, Ed.), *Numerical Mathematics, Singapore 1988*, ISNM 88, Birkhäuser, 1988.

[12] P.C. Hansen, *Regularization, GSVD and truncated GSVD*, BIT, to appear.

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

[14] P. Lintz, *Uncertainty in the solution of linear operator equations*, BIT 24 (1984), 92-101.

[15] A.K. Louis & F. Natterer, *Mathematical problems of computerized tomography*, Proc. IEEE 71 (1983), 379-389.

[16] I.P. Nedelkov, *Improper problems in computational physics*, Computer Physics Communications 4 (1972), 157-164.

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

[18] G.R. Richter, *Numerical solution of integral equations of the first kind with nonsmooth kernels*, SIAM J. Numer. Anal. 15 (1978), 511-522.

[19] E. Schock, *What are the proper condition numbers of discretized ill-posed problems?*, Lin. Alg. & Its Appl. 81 (1986), 129-136.

[20] W.M. Smart, *Stellar Dynamics*, Cambridge University Press (1938).

[21] A.N. Tikhonov, *Solution of incorrectly formulated problems and the regularizaton method*, Doklady Akad. Nauk SSSR 151 (1963), 501-504 = Soviet Math. 4 (1963), 1035-1038.

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

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

[24] J.M. Varah, *Pitfalls in the numerical solution of linear ill-posed problems*, SIAM J. Sci. Stat. Comput. 4 (1983), 164-176.

[25] G.M. Wing, *Condition numbers of matrices arising from the numerical solution of linear integral equations of the first kind*, J. Integral Equations 9 (Suppl.) (1985), 191-204.

[26] H. Zha & P.C. Hansen, *Regularization and the general Gauss-Markoff linear model*, in preparation.

Figure captions

Figure 1. Typical behavior of —— the regularized solution x_λ for a right-hand side b that satisfies the discrete Picard condition, and – – – the regularized solution $\tilde{x}_\lambda^{(e)}$ for a right-hand side e that doesn't.

Figure 2. Comparison of × TSVD solutions $\tilde{x}_k$ and —— regularized solutions $\tilde{x}_\lambda$ corresponding to a perturbed right-hand side $b + e$.

Figure 3. Plot of TSVD and regularized solutions for Phillips' example.

Figure 4. The computed singular values σ_i and coefficients $|u_i^T b| / \sigma_i$ for the second example with observed stellar parallaxes.

Figure 5. Plot of TSVD and regularized solutions for the second example with observed stellar parallaxes.

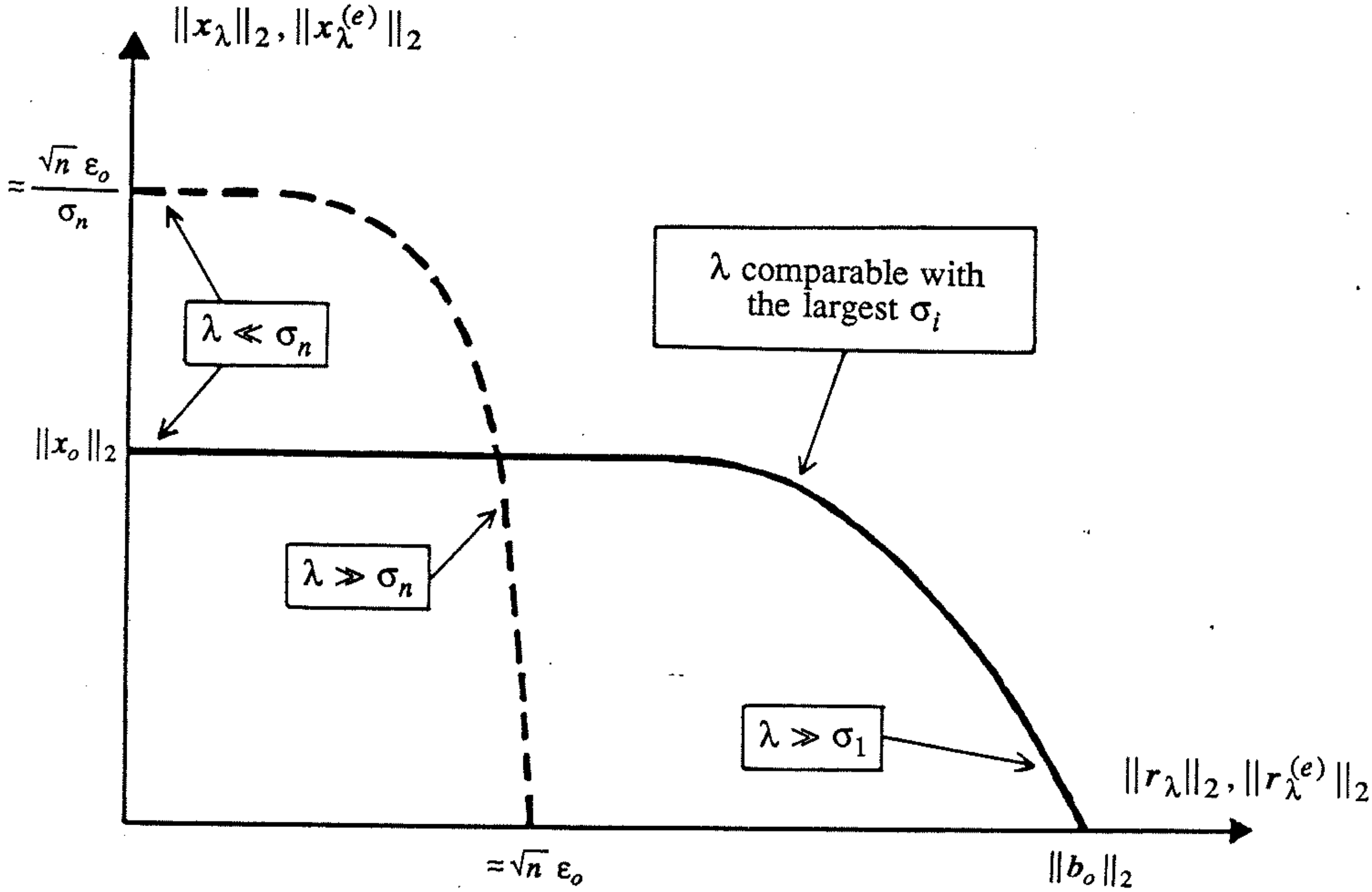
$\|x_\lambda\|_2, \|x_\lambda^{(e)}\|_2$
$\approx \frac{\sqrt{n}\,\varepsilon_o}{\sigma_n}$
$\lambda \ll \sigma_n$
$\|x_o\|_2$
$\lambda \gg \sigma_n$
λ comparable with the largest σ_i
$\lambda \gg \sigma_1$
$\|r_\lambda\|_2, \|r_\lambda^{(e)}\|_2$
$\approx \sqrt{n}\,\varepsilon_o$
$\|b_o\|_2$

Figure 1.

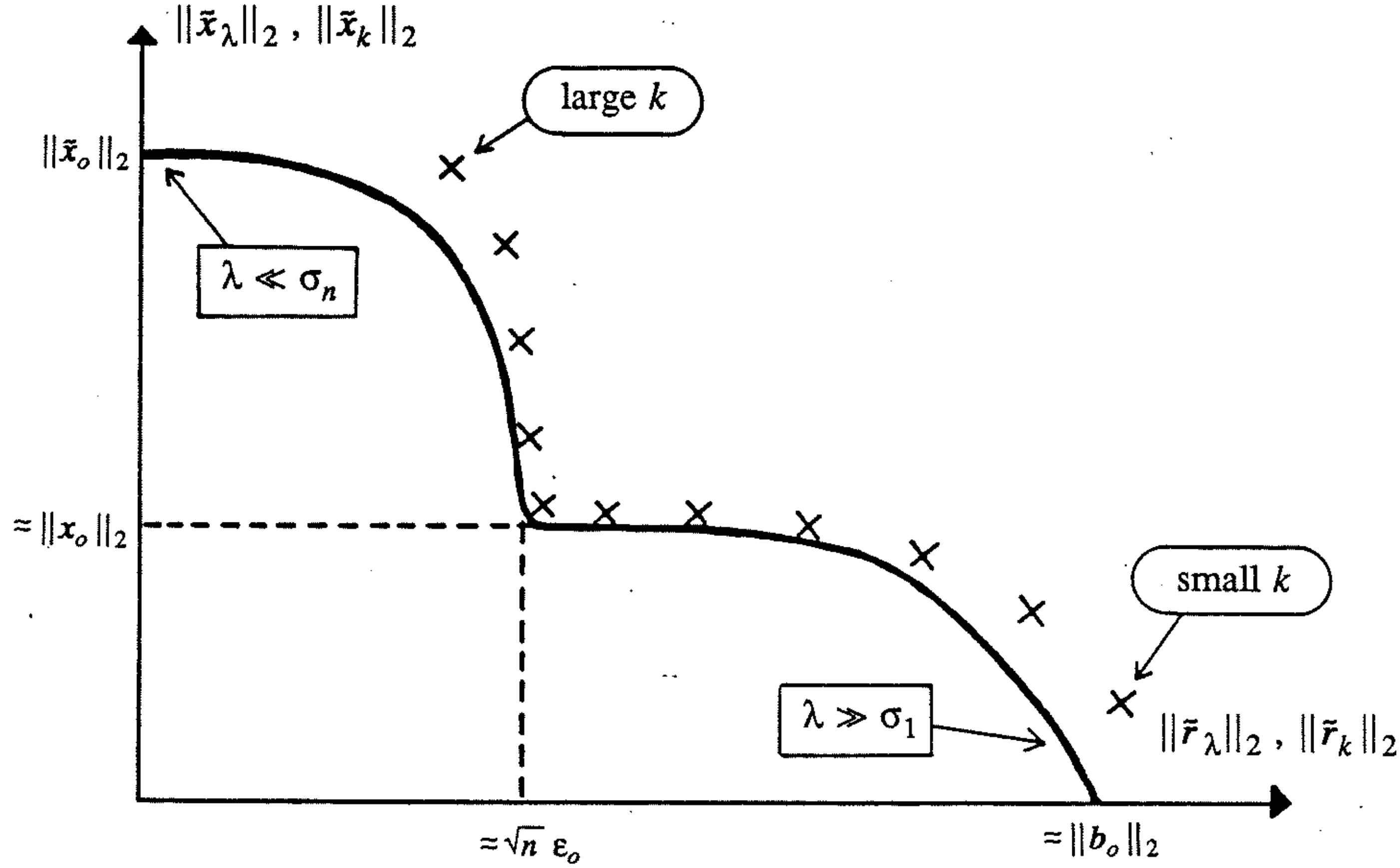
$\|\tilde{x}_\lambda\|_2$, $\|\tilde{x}_k\|_2$
large k
$\|\tilde{x}_o\|_2$
$\lambda \ll \sigma_n$
$\approx \|x_o\|_2$
small k
$\lambda \gg \sigma_1$
$\|\tilde{r}_\lambda\|_2$, $\|\tilde{r}_k\|_2$
$\approx \sqrt{n}\ \varepsilon_o$
$\approx \|b_o\|_2$

Figure 2

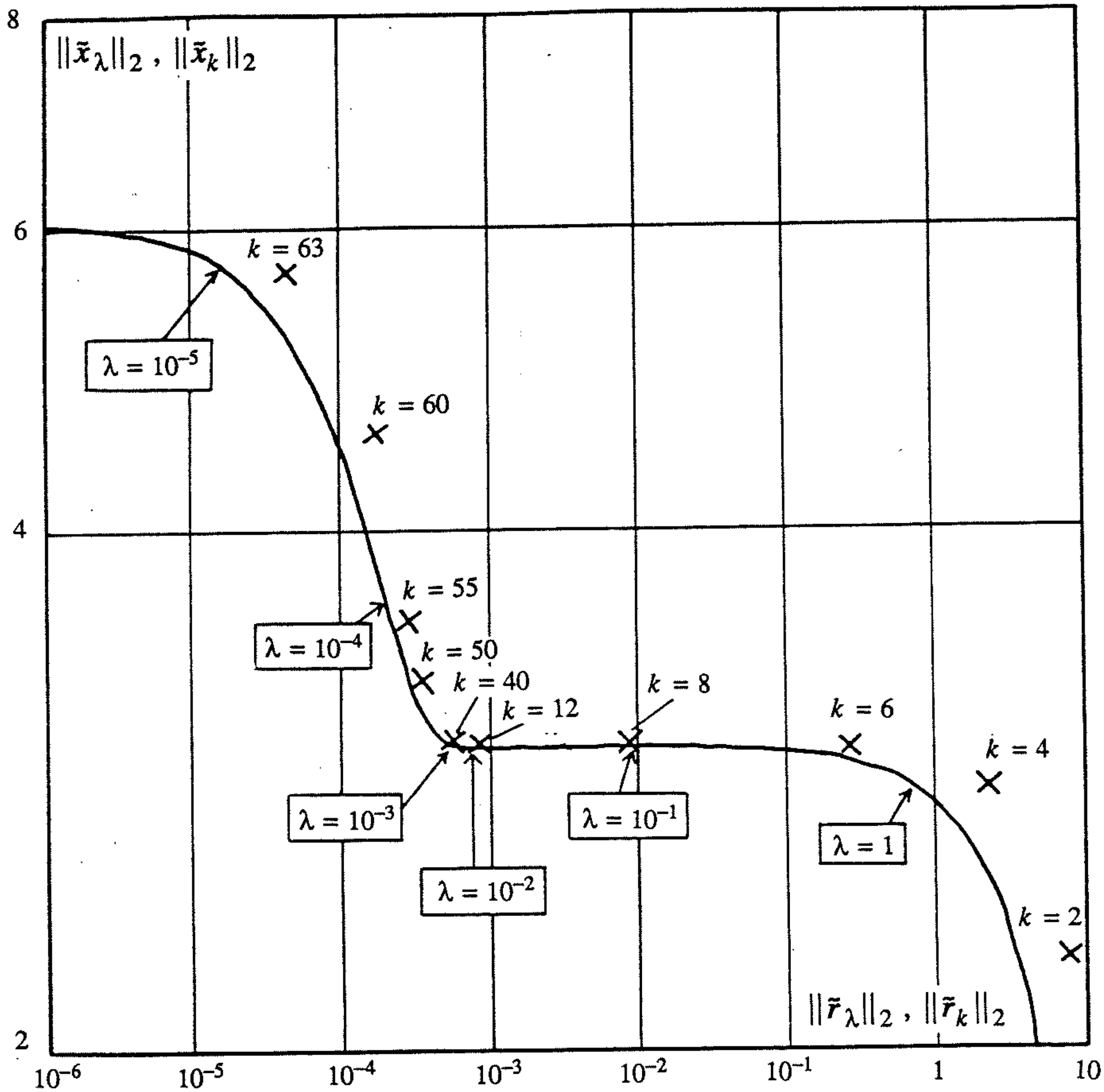
8
6
4
2
||x̃λ||2 , ||x̃k||2
||r̃λ||2 , ||r̃k||2
10−6
10−5
10−4
10−3
10−2
10−1
1
10
k = 63
k = 60
k = 55
k = 50
k = 40
k = 12
k = 8
k = 6
k = 4
k = 2
λ = 10−5
λ = 10−4
λ = 10−3
λ = 10−2
λ = 10−1
λ = 1

Figure 3

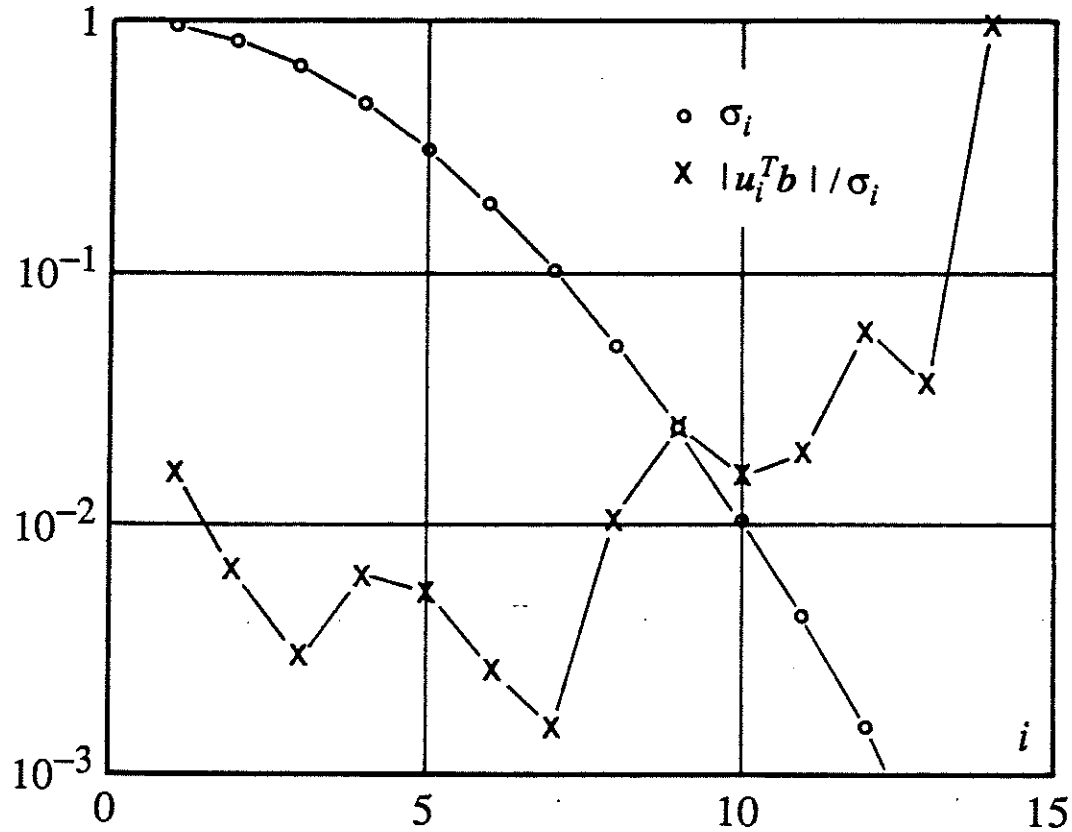

Figure 4

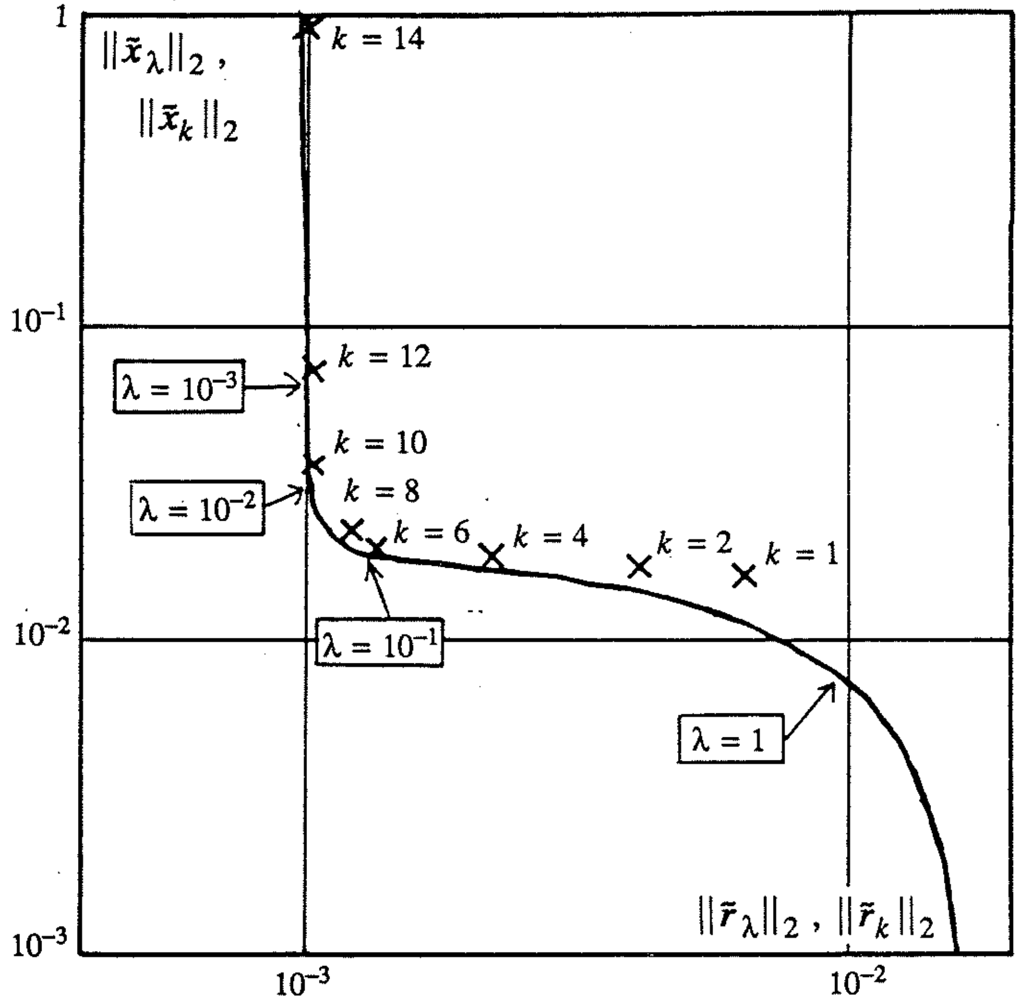
1
$\|\bar{x}_\lambda\|_2$,
$\|\bar{x}_k\|_2$
10^{-1}
10^{-2}
10^{-3}
$k = 14$
$k = 12$
$k = 10$
$k = 8$
$k = 6$
$k = 4$
$k = 2$
$k = 1$
$\lambda = 10^{-3}$
$\lambda = 10^{-2}$
$\lambda = 10^{-1}$
$\lambda = 1$
$\|\bar{r}_\lambda\|_2$, $\|\bar{r}_k\|_2$
10^{-3}
10^{-2}

Figure 5