UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

The Discrete Picard Condition for Discrete Ill-Posed Problems

Per Christian Hansen

July 1989

CAM Report 89-22

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

The Discrete Picard Condition for Discrete Ill-Posed Problems

Per Christian Hansen†

Dept. of Mathematics, UCLA, 405 Hilgard Ave., Los Angeles, CA 90024

July 1989

ABSTRACT

We investigate the convergence properties of regularized solutions to discrete ill-posed least squares problems. A necessary condition is that the Fourier coefficients of the right-hand side, when expressed in terms of the generalized SVD associated with the regularization problem, in average decay to zero faster than the generalized singular values. This is the discrete Picard condition. We illustrate the importance of this condition theoretically as well as experimentally.

AMS subject classification: 65F30, 65F20.

Key words: Ill-posed problems, Tikhonov regularization, discrete Picard condition, generalized SVD.

† The author is on leave from UNI•C, 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

By *discrete ill-posed problems*, we mean a particular class of discrete least squares problems

$$\min_x \|A\,x - b\|_2 , \qquad A \in \mathbb{R}^{m\times n} , \qquad m \ge n \tag{1.1}$$

where the singular values of the matrix A decay gradually to zero in such a fashion that A is very ill-conditioned. Often, due to rounding errors as well as errors in the data, such ill-conditioned matrices have full rank in a mathematical sense and the discrete least-squares problem (1.1) is therefore not ill-posed in the original sense due to Hadamard (see e.g. [12, Section 1.1]). However, we feel that it is still practical to use the terminology 'discrete ill-posed problems', partly because many of the difficulties of ill-posed problems carry over to the problem (1.1), and partly because (1.1) often arises when an underlying ill-posed problem — for example a Fredholm integral equation of the first kind — is discretized in order to compute a numerical solution.

For such problems, a variety of direct and iterative numerical regularization methods have been proposed, see e.g. [1,5,6,8,11,15,18,22,24,25,27] and the surveys in [2,3,9,12,19,20,28,29]. Many of these methods seek to either compute or approximate a certain regularized solution, namely the solution x_λ to the *discrete Tikhonov-regularization problem*

$$x_\lambda \;=\; \text{argmin}\,\{\|A\,x - b\|_2^2 + \lambda^2\,\|L\,x\|_2^2\} \tag{1.2}$$

where L typically is either the identity matrix or a well-conditioned discrete approximation to some derivative operator. Both the matrix L and the regularization parameter λ are used to control the smoothness of the regularized solution x_λ. An underlying assumption when using these methods is therefore that the exact solution, which one is trying to approximate by x_λ, is indeed smooth. Another assumption, which is equally important, is that the larger the singular values of A, the smoother the corresponding singular vectors (in the sense: less zero crossings). For a discussion of these aspects, see [28] and [15, Section 2].

There is, however, one more assumption which is not so well understood, and which bears a similarity with the Picard condition for ill-posed problems. Let $K(s,t) = \sum_{i=1}^{\infty} \sigma_i\, u_i(s)\, v_i(t)$ be the singular value expansion of the compact operator K, and let the right-hand side g be expressed as $g(s) = \sum_{i=1}^{\infty} \beta_i\, u_i(s)$. In order that the equation $K\,f = g$ have a square integrable least-squares solution f, it is necessary and sufficient that g satisfies the following condition [12, Theorem 1.2.6]:

The Picard condition (PC). *The right-hand side g in $K f = g$ satisfies the PC if*

$$\sum_{i=1}^{\infty} |(u_i , g)/\sigma_i|^2 < \infty \ , \quad \sigma_i \neq 0 \tag{1.3}$$

where (u_i , g) denotes the usual inner product between u_i and g.

Eq. (1.3) implies that from a certain point in the summation, the Fourier coefficients (u_i , g) must decay to zero faster than the σ_i. For the finite-dimensional discrete problem (1.1), the equivalent of Eq. (1.3) is always satisfied. Nevertheless, the rate between the decay of the singular values of A and the decay of the Fourier coefficients of the right-hand side b, when expressed in terms of the left singular vectors of A, still plays an important role for the success of discrete Tikhonov regularization. The purpose of this paper is to illustrate this phenomenon, via the introduction of the 'discrete Picard condition', and to show how this condition is used in practice. The work extends and further develops the authors work in [13,15,16]. The concept of a 'discrete Picard condition' was first discussed by Varah [28,29].

The paper is organized as follows. In Section 2 we introduce the generalized SVD, which we use throughout the paper to analyze Tikhonov regularization as well as a related regularization method, truncated GSVD. In sections 3 and 4, we investigate the conditions in which Tikhonov regularization and truncated GSVD will produce reasonable solutions. This analysis leads to the definition of the discrete Picard condition in Section 5, where we also briefly discuss how to test this condition numerically. Finally, in Section 6, we give two numerical examples.

2. Discrete Tikhonov regularization and generalized SVD

The most convenient tool for analysis of the discrete Tikhonov regularization-problem (1.2) is the *generalized SVD* (GSVD) of the matrix pair (A , L). The GSVD was introduced by Van Loan [26] and further generalized by Paige and Saunders [23]. Here, we use a slightly simpler formulation, which is sufficient for our analysis.

Theorem 1. *Let the matrix pair (A , L) satisfy*

$$A \in \mathbb{R}^{m \times n} \ , \quad L \in \mathbb{R}^{p \times n} \ , \quad m \geq n \geq p \ , \quad \mathrm{rank}(L) = p \ . \tag{2.1}$$

Then there exist matrices $U \in \mathbb{R}^{m \times n}$, $V \in \mathbb{R}^{p \times p}$ with $U^T U = I_n$, $V^T V = I_p$ and a nonsingular $X \in \mathbb{R}^{n \times n}$ such that

$$\begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} A \\ L \end{bmatrix} X^{-1} = \begin{bmatrix} \Sigma \\ M \end{bmatrix} = \begin{bmatrix} \Sigma_p & 0 \\ 0 & I_o \\ M_p & 0 \end{bmatrix} \begin{matrix} p \\ n-p \\ p \end{matrix} \tag{2.2}$$

$$\Sigma_p = \mathrm{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{p \times p}, \quad M_p = \mathrm{diag}(\mu_1, \ldots, \mu_p) \in \mathbb{R}^{p \times p} \tag{2.3}$$

where $\Sigma_p^2 + M_p^2 = I_p$ *and*

$$0 \le \sigma_1 \le \cdots \le \sigma_p \le 1, \qquad 1 \ge \mu_1 \ge \cdots \ge \mu_p > 0. \tag{2.4}$$

The generalized singular values of (A, L) *are defined as the ratios* $\gamma_i \equiv \sigma_i / \mu_i$, $i = 1, \ldots, p$.

Proof. See [23, Section 2]. □

Remark. As long as the matrix L is well-conditioned and its null-space is spanned by smooth vectors, it can be shown that the σ_i are closely related to the usual singular values ψ_j of A in the sense that

$$\|L^+\|_2^{-1} \le \psi_{n-i+1}/\sigma_i \le \|A\|_2 + \|L\|_2 \tag{2.5}$$

where L^+ is the pseudoinverse of L [16, Theorem 2.4 and Section 3].

Now, if we define $F_\lambda \equiv \mathrm{diag}(f_1, \ldots, f_p) \in \mathbb{R}^{p \times p}$ as a diagonal matrix with diagonal elements

$$f_i = \gamma_i^2 / (\gamma_i^2 + \lambda^2), \quad i = 1, \ldots, p \tag{2.6}$$

then it is easy to show that the regularized solution x_λ can be written as

$$x_\lambda = X \begin{bmatrix} F_\lambda \Sigma_p^+ & 0 \\ 0 & I_0 \end{bmatrix} U^T b = \sum_{i=1}^{p} f_i \frac{u_i^T b}{\sigma_i} x_i + \sum_{i=p+1}^{n} u_i^T b \; x_i . \tag{2.7}$$

Equations (2.6) and (2.7) show that λ basically 'filters out' the contributions to x_λ corresponding to small γ_i, and this illustrates how λ is used to control the sensitivity of x_λ to perturbations of A and b [17]. Also, since the oscillation property of the singular vectors of A carries over to the columns x_i of X (i.e., the larger the generalized singular values γ_i, the smoother the x_i) [16, Theorem 3.2], we see how we can use λ to control the smoothness of the solution. The 'cost' of this regularization is that we neglect a (hopefully) small part of the information in b and that x_λ is not an unbiased estimator.

In connection with our discussion of discrete Tikhonov regularization it is also natural to consider the *truncated GSVD* (TGSVD) solution introduced in [16], since this method is so closely related to both Tikhonov regularization and to the well-known truncated SVD method for regularization. Thus, define the TGSVD solution x_k as

$$x_k \equiv X \begin{bmatrix} \hat{\Sigma}_k^+ & 0 \\ 0 & I_o \end{bmatrix} U^T b, \quad \hat{\Sigma}_p^+ \equiv \mathrm{diag}(0, \ldots, 0, \sigma_{p-k+1}^{-1}, \ldots, \sigma_p^{-1}). \tag{2.8}$$

Since $\gamma_i = \sigma_i / \mu_i = \sigma_i / \sqrt{1 + \sigma_i^2}$, small γ_i correspond to small σ_i, so TGSVD simply corresponds to discarding the $p-k$ smallest γ_i of (A, L). The TGSVD solution x_k is also a regularized solution, and it is in fact very similar to x_λ in many respects, see [16] for more details. For $L = I_n$, TGSVD becomes the well-known truncated SVD method [13,15,28]. We include a discussion of x_k in this analysis in order to give an example of how the properties of the regularized solution x_λ carry over to other solutions that approximate x_λ, such as the TGSVD solution. The same analysis can, we believe, be carried our for any approximate regularized solution, once its expansion in terms of the GSVD is known. See for example the analysis of iterative methods in [27].

3. Regularization errors

One of the the most important topics in relation to any regularization method is the convergence analysis [12, Chapter 2]. For discrete problems, the central issue here is: how well does the regularized solution approximate the unknown, underlying exact solution. In this part of the analysis, one therefore ignores the errors in A and b and seeks to give a bound for the mere *regularization error*.

For both regularization methods considered here, Tikhonov regularization and TGSVD, the above-mentioned exact solution x_o is given by

$$x_o \;=\; X\,\Sigma^+ U^T b \;. \tag{3.1}$$

This is because x_o satisfies both $x_o = x_\lambda$ for $\lambda = 0$ (no Tikhonov regularization) and $x_o = x_k$ for $k = n$ (no truncation). Notice that if A does not have full rank, then the matrix $X\,\Sigma^+ U^T$ is, in general, different from the pseudoinverse A^+, cf. [7, Theorem 2.3]. However, in this case $x_o = X\,\Sigma^+ U^T b$ is still a member of the general solution to (1.1) given by $A^+ b + x^*$, where $A^+ b$ is the minimum norm least-squares solution to (1.1), and x^* is an arbitrary vector in the null space of A. Thus, it is always correct to compare the regularized solutions x_λ and x_k with x_o, and we therefore define $x_o - x_\lambda$ and $x_o - x_k$ as the *Tikhonov regularization error* and the *TGSVD regularization error*, respectively.

Our approach to analyzing these regularization errors is to derive bounds for the norms $\|L\,(x_o - x_\lambda)\|_2$ and $\|L\,(x_o - x_k)\|_2$ of the differences between the solutions (for $L = I_n$) or between the derivatives of the solutions (for $L \neq I_n$). Let U_p denote the matrix consisting of the first p columns of U. Then we can easily obtain the 'naive' bounds:

$$\|L\,(x_o - x_\lambda)\|_2 \;=\; \|M_p\,(I_p - F_\lambda)\,\Sigma_p^+ U_p^T b\,\|_2 \;\le\; \max\left\{\frac{\lambda^2}{\gamma_i\,(\gamma_i^2+\lambda^2)}\right\}\|b\|_2 \;\le\; \gamma_1^{-1}\,\|b\|_2$$

$$\|L\,(x_o - x_k)\|_2 \;=\; \|M_p\,(\Sigma_p^+ - \hat{\Sigma}_k^+)\,U_p^T b\,\|_2 \;\le\; \gamma_1^{-1}\,\|b\|_2\,.$$

However, these bounds certainly do not guarantee small regularization errors. Obviously, we must also incorporate into our analysis more information about the right-hand side b. I.e., we must analyze upper bounds of the form:

$$\|L(x_o - x_\lambda)\|_2 \le \sqrt{p}\, \|M_p (I_p - F_\lambda)\Sigma_p^+ U_p^T b\|_\infty = \sqrt{p} \max_{1\le i\le p} \left\{ \frac{\lambda^2}{\gamma_i^2+\lambda^2} \frac{|u_i^T b|}{\gamma_i} \right\} \tag{3.2a}$$

$$\|L(x_o - x_k)\|_2 \le \sqrt{p}\, \|M_p (\Sigma_p^+ - \hat{\Sigma}_k^+) U_p^T b\|_\infty = \sqrt{p} \max_{1\le i\le p-k} \left\{ \frac{|u_i^T b|}{\gamma_i} \right\}. \tag{3.2b}$$

In the light of these bounds, it is easy to see that the upper bounds for the regularization errors are related to the ratios between the Fourier coefficient $u_i^T b$ and the corresponding generalized singular values γ_i.

To emphasize this relationship, we use the same strategy as in [15,16] and assume a simple, but still realistic 'model' of the right-hand sides b as the typically appear in discrete ill-posed problems. In this 'model', we assume that the Fourier coefficients have the following simple form:

$$u_i^T b = \begin{cases} \gamma_i^\alpha & , \quad i = 1, \dots, p \\ \gamma_p^\alpha & , \quad i = p+1, \dots, m \end{cases} , \qquad \alpha \ge 0 . \tag{3.3}$$

Here, $\alpha \ge 0$ is a real parameter which controls the decay rate of the Fourier coefficients $u_i^T b$ relative to that of the generalized singular values γ_i. Then we have:

Theorem 2. *Let x_λ and x_k denote the regularized solutions (2.7) and (2.8), and let x_o denote the unregularized solution (3.1). Further, let the right-hand side b satisfy Eq. (3.3). Then the regularization errors satisfy*

$$\frac{\|L(x_o - x_\lambda)\|_2}{\|L x_o\|_2} \le \begin{cases} \sqrt{p} & , \quad 0 \le \alpha < 1 \\ \sqrt{p}\,(\lambda/\gamma_p)^{\alpha-1} & , \quad 1 \le \alpha < 3 \\ \sqrt{p}\,(\lambda/\gamma_p)^2 & , \quad 1 \le \alpha \end{cases} \tag{3.4a}$$

$$\frac{\|L(x_o - x_k)\|_2}{\|L x_o\|_2} \le \begin{cases} \sqrt{p} & , \quad 0 \le \alpha < 1 \\ \sqrt{p}\,(\gamma_{k-p+1}/\gamma_p)^{\alpha-1} & , \quad 1 \le \alpha . \end{cases} \tag{3.4b}$$

Proof. Inserting (3.3) into Eqs. (3.2a) and (3.2b), we are lead to the problem of determining

$$M_\lambda(\gamma_i) = \max_{1\le i\le p} \left\{ \frac{\lambda^2}{\gamma_i^2+\lambda^2} \gamma_i^{\alpha-1} \right\} \qquad \text{and} \qquad M_k(\gamma_i) = \max_{1\le i\le p-k} \{\gamma_i^{\alpha-1}\} .$$

For $0 \le \alpha < 1$, M_λ is a decreasing function, and $M_\lambda(\gamma_i) \le M_\lambda(\gamma_1) \le \gamma_1^{\alpha-1}$. Similarly, for $3 \le \alpha$, M_λ is an increasing function, and $M_\lambda(\gamma_i) \le M_\lambda(\gamma_p) \le \lambda^2 \gamma_p^{\alpha-3}$. And for $1 \le \alpha < 3$, $M_\lambda(\gamma)$ attains its maximum for $\gamma = \overline{\gamma} = \lambda^2(\alpha-1)(3-\gamma)$, and $M_\lambda(\gamma_i) \le M_\lambda(\overline{\gamma}) = \frac{1}{2}(\alpha-1)^{\frac{1}{2}(\alpha-1)}(3-\alpha)^{\frac{1}{2}(3-\alpha)}\lambda^{\alpha-1} \le \lambda^{\alpha-1}$. — Concerning M_k, it is easy to show that $M_k(\gamma_i) \le \gamma_1^{\alpha-1}$ for $0 \le \alpha < 1$ and $M_k(\gamma_i) \le \gamma_{p-k}^{\alpha-1}$ for $1 \le \alpha$. — Finally, we use the relation $L\,x_o = V\,M\,\Sigma^+ U^T b = V_p\,M_p\,\Sigma_p^+ U^T b$ to obtain the bound

$$\|L\,x_o\|_2 \;=\; \|M_p\,\Sigma_p^+ U_p^T b\|_2 \;\ge\; \|M_p\,\Sigma_p^+ U_p^T b\|_\infty \;=\; \max_{1 \le i \le p}\{\gamma_i^{\alpha-1}\} \;=\; \begin{cases} \gamma_1^{\alpha-1} & , \quad 0 \le \alpha < 1 \\ \gamma_p^{\alpha-1} & , \quad 1 \le \alpha\,. \end{cases}$$

Inserting all these bounds into Eqs. (3.2a) and (3.2b), we obtain (3.4a) and (3.4b). □

Remark. For $L = I_n \;\Rightarrow\; \gamma_{p-k+1} = \psi_k,\ \gamma_p = \psi_1$, Eqs. (3.4a) and (3.4b) are consistent with the results in [15, Theorem 3.1].

In any practical application of Tikhonov regularization and TGSVD, to obtain a reasonable 'filtering' of the small γ_i one always chooses $\lambda < \gamma_p$ or $k < p$. Theorem 2 then shows that in order to guarantee small regularization errors, α must be somewhat larger than 1, i.e. the absolute value of the Fourier coefficients, $|u_i^T b|$, must decay to zero *faster* than the generalized singular values γ_i. And the faster the decay, the better x_λ and x_k approximate x_o. This is, in fact, a very basic assumption, and it *must* be satisfied by the underlying, exact problem in order to ensure that Tikhonov regularization or TGSVD be able to produce useful approximate solutions.

4. Similarity of Tikhonov regularization and TGSVD

In this section we take a closer look at the similarity between Tikhonov regularization and TGSVD, and we investigate the conditions in which we can guarantee that x_k is indeed close to x_λ. For this purpose, it is convenient to introduce matrices A_λ^I and A_k^I such that x_λ and x_k can be written as $x_\lambda = A_\lambda^I\, b$ and $x_k = A_k^I\, b$. Eqs. (2.7) and (2.8) show that these matrices are uniquely determined by

$$A_\lambda^I \;=\; X \begin{bmatrix} F_\lambda \Sigma_p^+ & 0 \\ 0 & I_0 \end{bmatrix} U^T \quad \text{and} \quad A_k^I \;=\; X \begin{bmatrix} \hat{\Sigma}_k^+ & 0 \\ 0 & I_o \end{bmatrix} U^T\,. \tag{4.1}$$

Then the difference between the solutions x_λ and x_k and the difference between their residuals can be measured by the norms of $L\,A_\lambda^I - L\,A_k^I$ and $A\,A_\lambda^I - A\,A_k^I$:

Theorem 3. *Let the matrices A_λ^I and A_k^I be given by (4.1), and let $\omega_k = \gamma_{p-k}/\gamma_{p-k+1}$. Then for any $\lambda > 0$ and any positive $k < p$:*

$$\frac{\omega_k^{1/2}}{1+\omega_k^{1/2}} \;\le\; \min_\lambda \frac{\|(L\,A_\lambda^I - L\,A_k^I)\|_2}{\|L\,A_k^I\|_2} \;\le\; \frac{\omega_k^{1/2}}{1+\omega_k^{3/2}}\,. \tag{4.2a}$$

$$\min_{\lambda} \|A\,A_\lambda^I - A\,A_k^I\|_2 = \frac{\omega_k}{1+\omega_k} \,. \tag{4.2b}$$

These two minima are attained for $\lambda \approx (\gamma_{p-k+1}^3\,\gamma_{p-k})^{1/4}$ *and* $\lambda = (\gamma_{p-k+1}\,\gamma_{p-k})^{1/2}$*, respectively.*

Proof. Using Eqs. (2.6)-(2.8) and (4.1), we get $\|L\,A_\lambda^I - L\,A_k^I\|_2 = \|F_\lambda M_p\,\Sigma_p^+ - M_p\,\hat{\Sigma}_k^+\|_2 = \max\{f_1\gamma_1^{-1},...,f_{p-k}\,\gamma_{p-k}^{-1},(1-f_{p-k+1})\gamma_{p-k+1}^{-1},...,(1-f_p)\gamma_p^{-1}\}$ and $\|A\,A_\lambda^I - A\,A_k^I\|_2 = \|F_\lambda - \Sigma_p\,\hat{\Sigma}_k^+\|_2 = \max\{f_1,...,f_{p-k},(1-f_{p-k+1}),...,(1-f_p)\}$. Since f_i and $f_i\,\gamma_i^{-1}$ are increasing functions of γ_i, while $(1-f_i)$ and $(1-f_i)\gamma_i^{-1}$ are decreasing functions of γ_i, it follows that

$$\|L\,A_\lambda^I - L\,A_k^I\|_2 = \max\left\{\frac{\gamma_{p-k}}{\gamma_{p-k}^2+\lambda^2}\,,\,\frac{\lambda^2}{(\gamma_{p-k+1}^2+\lambda^2)\,\gamma_{p-k+1}}\right\} \tag{4.3a}$$

$$\|A\,A_\lambda^I - A\,A_k^I\|_2 = \max\left\{\frac{\gamma_{p-k}^2}{\gamma_{p-k}^2+\lambda^2}\,,\,\frac{\lambda^2}{\lambda^2+\gamma_{p-k+1}^2}\right\}. \tag{4.3b}$$

Equating the two terms in (4.3a), we are lead to the following expression:

$$\lambda^2 = \tfrac{1}{2}\left[\gamma_{p-k}\,(\gamma_{p-k+1}-\gamma_{p-k}) + \left[\gamma_{p-k}^2\,(\gamma_{p-k+1}-\gamma_{p-k})^2 + 4\gamma_{p-k+1}^3\,\gamma_{p-k}\right]^{1/2}\right] \approx \left[\gamma_{p-k+1}^3\,\gamma_{p-k}\right]^{1/2}.$$

Upper and lower bounds for $\|L\,A_\lambda^I - L\,A_k^I\|_2$ are then obtained by inserting this approximate value for λ into the two terms in (4.3a), yielding:

$$\gamma_{p-k+1}^{-1}\,\frac{\omega_k^{1/2}}{1+\omega_k^{1/2}} \le \|L\,A_\lambda^I - L\,A_k^I\|_2 \le \gamma_{p-k}\,\frac{\omega_k^{3/2}}{1+\omega_k^{3/2}} \,.$$

Finally, we divide these bounds by $\|L\,A_k^I\|_2 = \gamma_{p-k+1}^{-1}$ to obtain Eq. (4.2a). — Following the same technique, it is easy to show that the two terms in the expression (4.3b) are equal for $\lambda^2 = \gamma_{p-k+1}\,\gamma_{p-k}$, and then $\|A\,A_\lambda^I - A\,A_k^I\|_2 = \omega_k/(1+\omega_k)$. This proves Eq. (4.2b). □

Remark. For the special case $L = I_n \Rightarrow \omega_k = \psi_{k+1}/\psi_k$, Eqs. (4.2a) and (4.2b) agree with the results derived in [13, Theorem 5.2].

Theorem 3 guarantees similar results from Tikhonov regularization and TGSVD whenever ω_k is sufficiently small and λ is chosen somewhere between $(\gamma_{p-k+1}\,\gamma_{p-k})^{1/2}$ and $(\gamma_{p-k+1}^3\,\gamma_{p-k})^{1/4}$. If the matrix A has well-determined numerical rank, i.e. if there is a distinct gap in the singular value spectrum, and therefore also in spectrum of generalized singular values γ_i, then $\omega_k = \gamma_{p-k}/\gamma_{p-k+1}$ can always be made small by a proper choice of k. In this case, it is also *natural* to choose this k as the truncation parameter in the TGSVD method. Then Theorem 3 shows that there always exists a λ and a k such that

TGSVD produces results similar to those obtained by Tikhonov regularization. This is, in itself, an important result.

However, practical experience with the use of the TGSVD method suggests that is can also be used successfully as a regularization method when A has ill-determined numerical rank, i.e. when the singular values of A, and therefore also the generalized singular values of (A, L), decay gradually to zero without any particular gap in the spectrum (see e.g. [28,29]). In order to analyze this situation, we must again incorporate information about the right-hand side into our analysis, and we shall again use the simple 'model' from the previous section. We shall also assume that λ lies in the interval $[\gamma_{p-k}, \gamma_{p-k+1}]$, since we know from Theorem 3 that x_λ and x_k are most similar for such λ.

Theorem 4. *Let x_λ and x_k denote the regularized solutions (2.7) and (2.8), and let the right-hand side b satisfy Eq. (3.3). If $\gamma_{p-k} \le \lambda \le \gamma_{p-k+1}$, then*

$$\frac{\|L x_\lambda - L x_k)\|_2}{\|L x_k\|_2} \le \begin{cases} \sqrt{p}\ \omega_k^{\alpha-1} & , \quad 0 \le \alpha \le < 1 \\ \sqrt{p}\ (\gamma_{p-k+1}/\gamma_p)^{\alpha-1} & , \quad 1 \le \alpha < 3 \\ \sqrt{p}\ (\gamma_{p-k+1}/\gamma_p)^2 & , \quad \alpha \ge 3 \end{cases} \tag{4.4a}$$

$$\frac{\|A x_\lambda - A x_k\|_2}{\|b\|_2} \le \begin{cases} \sqrt{p}\ (\gamma_{p-k+1}/\gamma_p)^\alpha & , \quad 0 \le \alpha < 2 \\ \sqrt{p}\ (\gamma_{p-k+1}/\gamma_p)^2 & , \quad \alpha \ge 2\,. \end{cases} \tag{4.4b}$$

Proof. Using the same idea as in the proof for Theorem 3, we obtain

$$\|L x_\lambda - L x_k\|_2 \le \sqrt{p}\ \max\{f_1\gamma_1^{\alpha-1}, \ldots, f_{p-k}\gamma_{p-k}^{\alpha-1}, (1-f_{p-k+1})\gamma_{p-k+1}^{\alpha-1}, \ldots, (1-f_p)\gamma_p^{\alpha-1}\}$$

$$\|A x_\lambda - A x_k\|_2 \le \sqrt{p}\ \max\{f_1\gamma_1^{\alpha}, \ldots, f_{p-k}\gamma_{p-k}^{\alpha}, (1-f_{p-k+1})\gamma_{p-k+1}^{\alpha}, \ldots, (1-f_p)\gamma_p^{\alpha}\}\,.$$

Here, $f_i\gamma_i^{\alpha-1}$ and $f_i\gamma_i^\alpha$ are increasing functions of γ_i for $\gamma_i \le \lambda$, and since $\gamma_{p-k} \le \lambda \le \gamma_{p-k+1}$, we have $f_i\gamma_i^{\alpha-1} \le f_{p-k}\gamma_{p-k}^{\alpha-1} \le \gamma_{p-k}^{\alpha-1}$ and $f_i\gamma_i^\alpha \le f_{p-k}\gamma_{p-k}^\alpha \le \gamma_{p-k}^\alpha$. We also have $(1-f_i)\gamma_i^{\alpha-1} \le \gamma_{p-k+1}^2\gamma_i^{\alpha-3}$ which is bounded above by $\gamma_{p-k+1}^{\alpha-1}$ for $0 \le \alpha \le 3$ and by $\gamma_{p-k+1}^2\gamma_p^{\alpha-3}$ for $3 \le \alpha$. Similarly, $(1-f_i)\gamma_i^\alpha \le \gamma_{p-k+1}^2\gamma_i^{\alpha-2}$, which is bounded above by γ_{p-k+1}^α for $0 \le \alpha \le 2$ and by $\gamma_{p-k+1}^2\gamma_p^{\alpha-2}$ for $2 \le \alpha$. Hence, we obtain the bounds

$$\|L x_\lambda - L x_k\|_2 \le \begin{cases} \sqrt{p}\ \gamma_{p-k}^{\alpha-1} & , \ 0 \le \alpha < 1 \\ \sqrt{p}\ \gamma_{p-k+1}^{\alpha-1} & , \ 1 \le \alpha < 3 \\ \sqrt{p}\ \gamma_{p-k+1}^2\gamma_p^{\alpha-3} & , \ 3 \le \alpha \end{cases} \quad \|A x_\lambda - A x_k\|_2 \le \begin{cases} \sqrt{p}\ \gamma_{p-k+1}^\alpha & , \ 0 \le \alpha < 2 \\ \sqrt{p}\ \gamma_{p-k+1}^2\gamma_p^{\alpha-2} & , \ 2 \le \alpha\,. \end{cases}$$

To obtain the results in (4.4a) and (4.4b), we combine these bounds with

$$\|L\,x_k\|_2 \;\geq\; \|M_p\,\hat{\Sigma}_k^+ U_p^T b\|_\infty \;=\; \max\{\gamma_{p-k+1}^{\alpha-1},\ldots,\gamma_p^{\alpha-1}\} \;\geq\; \begin{cases} \gamma_{p-k+1}^{\alpha-1} & , \quad 0 \leq \alpha < 1 \\ \gamma_p^2\,\gamma_{p-k}^{\alpha-3} & , \quad 1 \leq \alpha \end{cases}$$

and with $\|b\|_2 \geq \|U^T b\|_\infty = \gamma_p^\alpha \leq \gamma_p^2 \gamma_{p-k}^{\alpha-2}$. □

Remark. When $L = I_n \Rightarrow \gamma_{p-k+1} = \psi_k,\ \gamma_p = \psi_1$, Eqs. (4.4a) and (4.4b) agree with the results in [15, Theorem 3.2].

Theorem 4 extends the results in Theorem 3: it shows that if α is somewhat larger than one, such that the Fourier coefficients $u_i^T b$ decay to zero faster than the γ_i, then there exist λ and k such that the regularized solutions x_λ and x_k are very similar, even if there is no particular gap in the singular value spectrum of A. In this connection, it is interesting to notice that practical choices of λ and k, based on e.g. *generalized cross-validation* (GCV) [10], always produces λ and k satisfying $\gamma_{p-k} \leq \lambda \leq \gamma_{p-k+1}$, see [13, Section 5]. The important conclusion is therefore that whenever Tikhonov regularization produces a satisfactory regularized solution x_λ, the TGSVD is also guaranteed to produce a satisfactory solution x_k which is very similar to x_λ.

5. The discrete Picard condition

As we have seen in the previous sections, the decay rate of the Fourier coefficients indeed plays a central role in connection with discrete ill-posed problems. The key result is that the regularized solution x_λ or x_k is only guaranteed to approximate the exact solution x_o if the Fourier coefficients $|u_i^T b|$ decay to zero faster than the generalized singular values γ_i. Of course, the decay of the Fourier coefficients needs not be monotonic, as long as they *in average* decay to zero faster than the γ_i.

There is an important exception to this requirement to the Fourier coefficients. If some singular values ψ_j of A are *numerically zero* (i.e., smaller than some threshold reflecting the errors in A), then the corresponding generalized singular values γ_i are also small due to (2.5), and the decay of the corresponding $|u_i^T b|$ is not important. Instead, the size of these $|u_i^T b|$ largely determines the norm of the residual and therefore, in turn, signal whether the problem (1.1) is consistent or not. This has nothing to do with the existence of a smooth solution, and it is therefore important to consider the decay of the Fourier coefficients corresponding to numerically nonzero generalized singular values (or, if $L = I_n$, singular values) only.

This discussion leads to the following definition of a *discrete Picard condition* for discrete ill-posed problems:

The discrete Picard condition (DPC). *Let* b *denote an unperturbed right-hand side in (1.2). Then* b *satisfies the DPC if, for all numerically nonzero generalized singular values* γ_i, *the Fourier coefficients* $|u_i^T b|$ *in average decay to zero faster than the* γ_i.

When solving real-world problems, where the right-hand side (and sometimes also the matrix) are contaminated with measurement errors, approximation errors, and rounding errors, then the given, perturbed problem rarely satisfies the DPC. However, if the underlying exact problem satisfies the DPC, then by a proper choice of λ or k one can often make the regularized problem satisfy the DPC. I.e., one can regard regularization as a method to derive from the given ill-posed problem a related problem that satisfies the DPC and therefore has a regularized solution that approximates the exact, unknown solution.

As a typical example of this situation, let the problem (1.1) be derived from a first kind Fredholm integral equation satisfying the Picard condition. Then, ideally, due to the strong connection between the singular value expansion of the kernel and the SVD of the matrix A [14], the DPC is also satisfied. However, due to data errors as well as approximation errors in setting up the discrete problem, all the $|u_i^T b|$ do not satisfy the DPC. Instead, they typically 'roll off' gradually until they reach an almost constant level, determined by the errors. By means of a proper choice of the Tikhonov regularization parameter λ one can, however, guarantee that the Fourier coefficients for the regularized problem, namely $|f_i \, u_i^T b|$, satisfy the DPC. Similarly, for TGSVD, one can choose a truncation parameter k such that the $|u_i^T b|$ satisfy the DPC for $i > p-k$. We give examples of this behavior in Section 6.

If, on the other hand, the underlying problem does not satisfy the DPC (or even the PC), then it is generally not possible to compute a satisfactory solution by means of Tikhonov regularization or any related method. See [4] for an example of this situation.

Having introduced the discrete Picard condition as defined above, a natural question is: how does one check numerically whether the DPC is satisfied? Of course, a visual inspection of a plot of the Fourier coefficients $|u_i^T b|$ and the generalized singular values γ_i will often reveal this and, at the same time, guide the user in choosing a suitable k or λ. But we believe that an automatic check for satisfaction of the DPC may also be required, partly because the amount of data may be large, and partly because one might want a more quantitative test than just a visual inspection (as is the case for the GCV method for choosing the optimal regularization parameters λ and k).

From the discussion leading to the DPC, it is evident that satisfaction of the DPC is a 'local' phenomenon, taking place only for the larger generalized singular values γ_i. Hence, the check for satisfaction of the DPC should also be based on the use of 'local' information in the sequences $|u_i^T b|$ and

γ_i, only. We could, for example, fit cubic splines to the $|u_i^T b|$ and the γ_i and then check the relative decay of these splines. However, a much simpler and easy-to-use approach seems to be sufficient.

Since we are interested in information about the decay of the data, it is the *ratios* of nearby coefficients $|u_i^T b|$ and γ_i — rather than their absolute values — that is important. Therefore, we propose to base the numerical check for satisfaction of the DPC on the *moving geometric mean:*

$$\rho_i \equiv \gamma_i^{-1} \left[\prod_{j=i-q}^{i+q} |u_j^T b| \right]^{\frac{1}{2q+1}} , \qquad i = q+1, \ldots, n-q \tag{5.1}$$

where q is a small integer, thus ensuring the locality of the ρ_i. Note that ρ_i should only be computed for numerically nonzero γ_i, and that special care should also be taken if some of the $|u_i^T b|$ are numerically zero. Based on our experiments, we find that q equal to 1, 2 or 3 gives good results, and we will say that the DPC is satisfied when the all ρ_i defined by (5.1), corresponding to numerically nonzero $|u_i^T b|$ and γ_i, decay monotonically to zero.

6. Numerical examples

In this section we illustrate by two numerical examples the important role the discrete Picard condition plays in the analysis of discrete ill-posed problems. Both examples are obtained from discretizations of Fredholm integral equations of the first kind:

$$\int_a^b K(s,x)\,dx = g(s) , \qquad c \le s \le d . \tag{6.1}$$

Our first example is the classical integral equation deviced by Phillips [24] with $[a,b] = [c,d] = [-6,6]$, and K and g given by:

$$K(s,x) = f(s-x) \equiv \begin{cases} 1 + \cos[(s-x)\pi/3] & , \quad |s-x| \le 3 \\ 0 & , \quad |s-x| > 3 \end{cases} \tag{6.2a}$$

$$g(s) = (6-|s|) \left[1 + \tfrac{1}{2} \cos \frac{\pi s}{3} \right] + \frac{9}{2\pi} \sin \frac{\pi |s|}{3} . \tag{6.2b}$$

This equation satisfies the Picard condition, and the square integrable solution is simply $f(x)$ as given in (6.2a). We discretized the integral equation using the trapezoidal quadrature rule as described in [21] with $m = 78$ and $n = 49$, and as regularization matrix L we chose an approximation to the second derivative operator,

$$L = \begin{bmatrix} -1 & 2 & -1 & & & \mathcal{O} \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & \mathcal{O} & & & & \\ & & & -1 & 2 & -1 \end{bmatrix} \in \mathbb{R}^{p \times n}, \qquad p = n - 2. \tag{6.3}$$

In order to simulate measurement errors in the right-hand side $\bar{b}$, we then added to b a random perturbation vector e with elements from a normal distribution with zero mean and standard deviation 10^{-5}.

The perturbed Fourier coefficients $|u_i^T \bar{b}|$ and the generalized singular values γ_i are shown in Fig. 1. Notice the 'reverse' ordering of the γ_i as compared to the usual singular values ψ_i. All the singular values of the kernel K in (6.2b) are nonzero, but due to the particular discretization method the matrix A has 7 numerically zero singular values ψ_i. Hence, there are also 7 numerically zero generalized singular values γ_i of (A, L), while the rest of the γ_i decay monotonically. The corresponding Fourier coefficients also tend to decay until they 'level off' at about 10^{-5}, which is the 'noise level' caused by our random errors e. The particular Fourier coefficients corresponding to the numerically zero γ_i are of the same size as the 'noise level', and the problem is therefore consistent within the 'noise'.

For this particular problem, all the Fourier coefficients $|u_i^T \bar{b}|$ with *even* index i are actually numerically zero (i.e., of the same size as the 'noise level'). As mentioned in Section 5, these numerically zero $|u_i^T \bar{b}|$ should not be included in the analysis of the DPC, and we therefore only computed the ρ_i for *odd* values of i. These ρ_i, computed with $q = 1$, are also shown in Fig. 1. We see that for $i = 45, 43, 41, 39, 37, 35$ the ρ_i decay monotonically, and for $i < 35$ the ρ_i start growing again. From such a plot, we conclude that the underlying, exact problem (with $e = 0$) seems to satisfy the discrete Picard condition. We also see that the perturbed problem satisfies the DPC if it is regularized in such a way that the terms $(u_i^T \bar{b} / \gamma_i) x_i$ are damped for $i < 35$. These observations perfectly agree with the appearance of the TGSVD solutions x_k (not shown here): for small k, x_k is dominated by approximation errors, while for k greater than about 15 the solution x_k is severely distorted by the influence of the errors. For k in the range 9 to 14, x_k is acceptably smooth and resembles the true solution.

The second example is a model of the transient transport across the blood-retina barrier in the human eye. A simple version of this model, assuming the eye to be a sphere, was refined to allow for a more realistic geometry of the eye where the front half deviates from spherical form. The kernel and right-hand side of the integral equation (6.1) are then given by:

$$K(s,x) = 2\pi \sin x \, D(x) \left[D(x)^2 + \left[\frac{d\,D(x)}{d\,x} \right]^2 \right]^{1/2} E(s^2 + D(x)^2 - 2\,s\,D(x) \cos x) \tag{6.4a}$$

$$g(s) = F(\gamma s + (\gamma - 1)a_o), \qquad \gamma = \tfrac{1}{2}(3+\delta)/(1+\delta) \tag{6.4b}$$

where $[a,b] = [0,\pi]$, $[c,d] = [-a_o, \delta a_o]$, and

$$D(x) = \begin{cases} \tfrac{1}{2}(\delta-1)a_o \cos 3x + \tfrac{1}{2}(\delta+1)a_o & , \quad 0 \le x < \pi/3 \\ a_o & , \quad \pi/3 \le x \le \pi \end{cases} \tag{6.5c}$$

$$E(t) = \frac{e^{-s_o t^{1/2}}}{t^{1/2}}, \qquad F(t) = \frac{e^{(t-a_o)s_o} - e^{-(t+a_o)s_o}}{t}. \tag{6.5d}$$

The parameters were $a_o = 1.2\,\mathrm{cm}$ (radius of the spherical back half of the eye), $\delta = 0.9$ (deformation from spherical form of the front half of the eye), and $s_o = 0.316$ (time constant for the diffusion through the blood-retina barrier). We discretized the integral equation by means of the moment method using piecewise constant approximations (see e.g. [14] for more details), and we used $L = I_n$. This leads us to considering the usual singular values ψ_i, the Fourier coefficients $|u_i^T \bar{b}|$ (where u_i are the usual left singular vectors of A), and the moving geometric mean $\rho_i = \psi_i^{-1} \left(\sum_{j=i-q}^{i+q} |u_j^T \bar{b}| \right)^{1/(2q+1)}$.

When the order n of the matrix is smaller than about 100, the DPC does not seem to be satisfied, because all the Fourier coefficients decay to zero slower than the singular values. However, if n is greater than 100, then the DPC is satisfied for $i < i_o$ (i.e., for the larger singular values), and the range i_o increases with the order n. This behavior of the i_o implies that the growth of the ρ_i for $i > i_o$ is primarily due to the approximation errors caused by the 'rough' piecewise constant approximations. A typical plot of the first 30 coefficients for the case $n = 256$ is show in Fig. 3, using $q = 3$. From these results, we can conclude that the underlying, exact problem seems to satisfy the DPC, which means that the integral equation model is satisfactory. Of course, use of more sophisticated approximation functions would increase the quality of the discrete solution.

Acknowledgements

I would like to thank Prof. Dianne P. O'Leary for discussions about the discrete Picard condition that highly influenced this presentation, and Prof. Tony F. Chan for providing very nice working conditions during my visit to UCLA.

References

[1] Å. Björck, *A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations*, BIT **28** (1988), 659-670.

[2] Å. Björck & L. Eldén, *Methods in numerical algebra for ill-posed problems*, Report LiTH-MAT-R33-1979, Dept. of Mathematics, Linköping University, 1979.

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

[4] 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*, pp. 92-101, Birkhäuser, 1980.

[5] M. Eiermann, I. Marek & W. Niethammer, *On the solution of singular linear systems of algebraic equations by semiiterative methods*, Numer. Math. **53** (1988), 265-283.

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

[7] L. Eldén, *A weighted pseudoinverse, generalized singular values, and constrained least squares problems*, BIT **22** (1982), 487-502.

[8] L. Eldén, *An algorithm for the regularization of ill-conditioned, banded least squares problems*, SIAM J. Sci. Stat. Comput. **5** (1984), 237-254.

[9] H.W. Engl & C.W. Groetsch (Eds.), *Inverse and Ill-Posed Problems*, Academic Press, 1987.

[10] 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.

[11] J. Graves & P.M. Prenter, *Numerical iterative filters applied to first kind Fredholm integral equations*, Numer. Math. **30** (1978), 281-299.

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

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

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

[15] P.C. Hansen, *Truncated SVD solutions to discrete ill-posed problems with ill-determined numerical rank*, SIAM J. Sci. Stat. Comput., to appear.

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

[17] P.C. Hansen, *Perturbation bounds for discrete Tikhonov regularization*, Inverse Problems, to appear.

[18] R.J. Hanson, *A numerical method for solving Fredholm integral equations of the first kind using singular values*, SIAM J. Numer. Anal. **8** (1972), 883-890.

[19] F. Natterer, *The Mathematics of Computerized Tomography*, Wiley, 1986.

[20] F. Natterer, *Numerical treatment of ill-posed problems;* in G. Talenti (Ed.), *Inverse Problems*, pp. 142-167, Lecture Notes in Mathematics 1225, Springer, 1986.

[21] D.P. O'Leary & B.W. Rust, *Confidence intervals for inequality-constrained least squares problems, with applications to ill-posed problems*, SIAM J. Sci. Stat. Comput. **7** (1986), 473-489.

[22] D.P. O'Leary & J.A. Simmons, *A bidiagonalization-regularization procedure for large scale discretizations of ill-posed problems*, SIAM J. Sci. Stat. Comput. **2** (1981), 474-489.

[23] C.C. Paige & M. Saunders, *Towards a generalized singular value decomposition*, SIAM J. Numer. Anal. **18** (1981), 398-405.

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

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

[26] C.F. Van Loan, *Generalizing the singular value decomposition*, SIAM J. Numer. Anal. **13** (1976), 76-83.

[27] A. van der Sluis & H.A. van der Vorst, *SIRT and CG type methods for the iterative solution of sparse linear least squares problems*, Report 88-43, Faculty of Technical Mathematics and Informatics, Delft University of Technology, 1988.

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

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

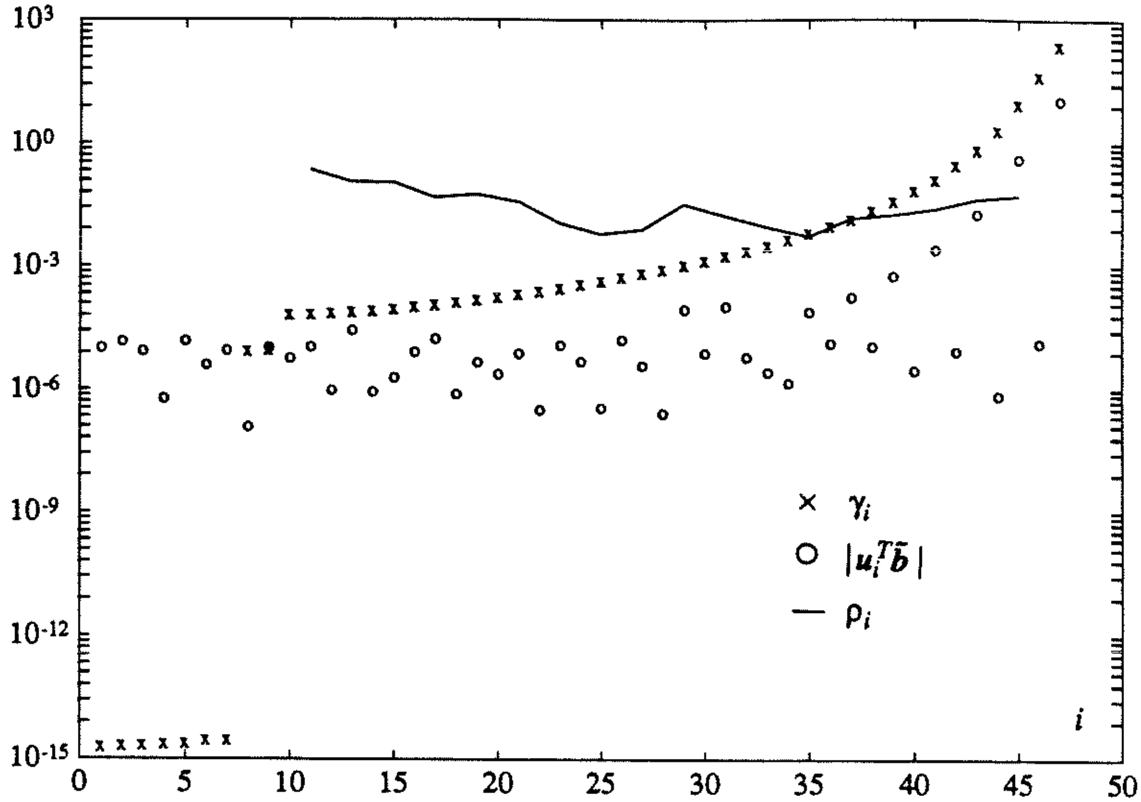

Figure 1. The generalized singular values (crosses), the Fourier coefficients $|u_i^T \tilde{b}|$ (circles), and the means ρ_i (solid line) for example one. The DPC is satisfied for $i = 45, 43, 41, 39, 37, 35$.

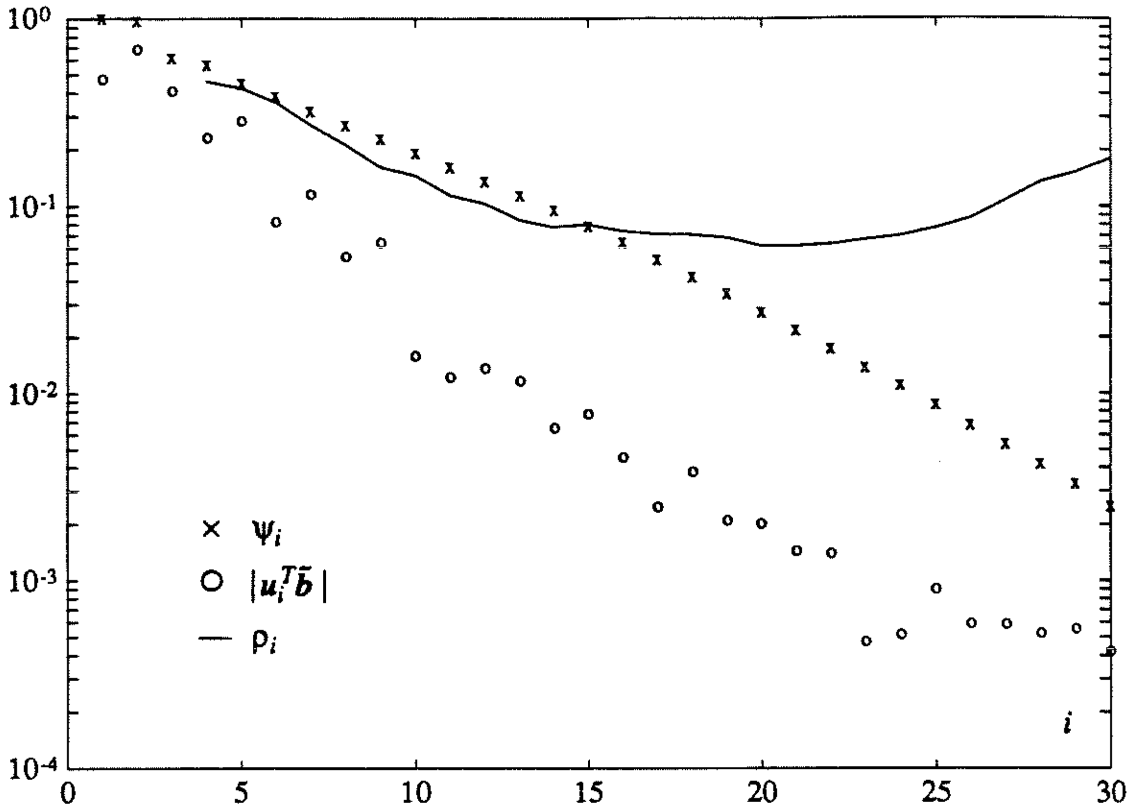

Figure 2. The singular values ψ_i (crosses), the Fourier coefficients $|u_i^T\tilde{b}|$ (circles), and the means ρ_i (solid line) for example one. The DPC is satisfied for $i < i_o = 20$.