University of Maryland College Park

Institute for Advanced Computer Studies TR–91–98-62
Department of Computer Science TR–3949

Three Results on Iterative Regularization*

Misha Kilmer† and G. W. Stewart‡

October 1998

ABSTRACT

In this paper we present three theorems which give insight into the regularizing properties of MINRES. While our theory does not completely characterize the regularizing behavior of the algorithm, it provides a partial explanation of the observed behavior of the method. Unlike traditional attempts to explain the regularizing properties of Krylov subspace methods, our approach focuses on convergence properties of the residual rather than on convergence analysis of the harmonic Ritz values. The import of our analysis is illustrated by two examples. In particular, our theoretical and numerical results support the following important observation: in some circumstances the dimension of the optimal Krylov subspace can be much smaller than the number of the components of the truncated spectral solution that must be computed to attain comparable accuracy.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports` or on the web at `http://www.cs.umd.edu/ stewart/`.

†ECE Department, Northeastern University, Boston, MA 02115. This work was supported in part by the Army Research Office under Grant DAAG55-97-1-0013.

‡Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. This work was supported by the National Science Foundation under grant CCR-95-03126.

Three Results on Iterative Regularization

Misha Kilmer
G. W. Stewart

ABSTRACT

In this paper we present three theorems which give insight into the regularizing properties of MINRES. While our theory does not completely characterize the regularizing behavior of the algorithm, it provides a partial explanation of the observed behavior of the method. Unlike traditional attempts to explain the regularizing properties of Krylov subspace methods, our approach focuses on convergence properties of the residual rather than on convergence analysis of the harmonic Ritz values. The import of our analysis is illustrated by two examples. In particular, our theoretical and numerical results support the following important observation: in some circumstances the dimension of the optimal Krylov subspace can be much smaller than the number of the components of the truncated spectral solution that must be computed to attain comparable accuracy.

1. Introduction

The discretization of a compact self-adjoint operator equation with errors in the right hand side results in a linear system whose exact solution bears no relation to the solution of the original error-free system. Specifically, in the spectral domain — that is, in the coordinate system of the eigenvectors of the matrix — the exact linear system can be written in the form

$$\Lambda x = b,$$

where

$$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \qquad 1 = \lambda_1 > \cdots > \lambda_n > 0.$$

Note that we assume that the system is positive definite and has been scaled so that $\lambda_1 = 1$. To the extent that the quantities in this equation track the corresponding quantities in the original operator, they have the following properties.

1. If the operator is smooth, the eigenvalues λ_i of the discrete operator will eventually decay rapidly to zero.

2. If the solution of the original problem is square integrable, then the components ξ_i of x decay to zero.

3. Since the components β_i of the right-hand side satisfy $\beta_i = \xi_i \lambda_i$, they decay faster than either the components of the solution or the eigenvalues of the operator.

The system with error is

$$\Lambda \tilde{x} = b + e \equiv \tilde{b}.$$

If we model the error as white noise, then we can append the following condition to the above list.

4. The components ϵ_i of e are random variables with mean zero and standard deviation ϵ.

The components of the solution for the perturbed right-hand side are

$$\tilde{\xi}_i = \frac{\beta_i + \epsilon_i}{\lambda_i}. \tag{1.1}$$

Since the β_i decay rapidly, they soon fall below the error level. The subsequent components of the solution are effectively random variables with standard deviation ϵ/λ_i. The division by λ_i magnifies the originally small error so that it dominates the solution.

As long as the β_i are larger than the errors ϵ_i the $\tilde{\xi}_i$ contain useful information. The process of extracting this information is called *regularization.* There are many regularization schemes. A natural one is to stop computing components of x when the β_i get near the error level. We will call this procedure *truncated spectral regularization.*[1]

In this paper we will be concerned with regularization based on the iterative method MINRES [8]. The method can be described briefly as follows. Let

$$K_k = (\tilde{b}\ \Lambda\tilde{b}\ \cdots\ \Lambda^{k-1}\tilde{b})$$

be the kth *Krylov matrix*. Then we seek an approximation to x in the form

$$x_k = K_k a_k,$$

where a_k is determined so that

$$\rho_k^2 \equiv \|\tilde{b} - \Lambda x_k\|^2 = \min.$$

Here $\|\cdot\|$ denotes the usual Euclidean norm.

The residual $r_k = \tilde{b} - \Lambda x_k$ has a alternate expression that is used in the analysis of the algorithm. Given a vector $a_k = (\alpha_1, \ldots, \alpha_k)^{\mathrm{T}}$, let a polynomial p be defined by

$$p(t) = 1 - \alpha_1 t - \alpha_2 t^2 - \cdots - \alpha_k t^k.$$

[1] In the nonsymmetric case this process is also known as truncated SVD regularization (see [6] and the references therein).

Then the residual $r_k = \tilde{b} - \Lambda K_k a_k$ can be written in the form $r_k = p_k(\Lambda)\tilde{b}$, where p_k is the polynomial associated with a_k. Since MINRES minimizes $\|r_k\|$, we have

$$\rho_k \equiv \|r_k\| = \min_{\substack{\deg(p_k)=k \\ p_k(0)=1}} \|p(\Lambda)\tilde{b}\|. \tag{1.2}$$

Because the ith component of the residual is the value of a polynomial at λ_i, we will often call it the value of the residual at λ_i.

It has been widely observed that up to a certain index k_0 the MINRES solutions x_k are increasingly accurate approximates to the exact solution x, after which their accuracy rapidly deteriorates. Thus if we can determine k_0 (a difficult problem), MINRES can be used as a regularization method. What seems less well known is that in some circumstances the critical index k_0 can be quite small — much smaller than the number of components of the truncated spectral solution that must be computed to attain the same accuracy.[2]

In this paper we give three theorems that help explain the regularizing properties of MINRES. The phrase "help explain" is a deliberate warning to the reader not to expect a complete analysis. The theorems do not say that MINRES has to behave in a particular way. Nonetheless, they give considerable insight into the method.

There is a vast body of literature on regularization, much of which is devoted to determining the asymptotic behavior of the error in the optimally regularized solution as the error in the right hand side approaches zero (this literature has been admirably surveyed by Engl, Hanke, and Neubauer [4]). There is also a body of literature for well-posed problems in which the asymptotic convergence of Krylov methods like MINRES are analyzed (see, for example, the recent book by Greenbaum [5]). Our approach, which consists in leaving the error in b fixed and determining how the solution behaves as the regularization parameter varies, is essentially nonasymptotic and fits in neither of these categories. On the other hand, our work is closely related to analyses that attempt to find the optimal value of the regularization parameter (for a survey of this literature see Hansen [7]).

In the next section we give a toy example that will be used to illustrate the subsequent results. We will also describe in a general way what is happening. The next three sections are devoted to establishing our three results. In Section 6 we apply our results to another example. The paper concludes with a discussion of the results and suggestions for future work.

A little informal terminology will help in our discussions. We will call the part of the spectrum for which the $\tilde{\beta}_i$ are little affected by the error the *initial part* of the spectrum. We will call the part where they are fully contaminated the *terminal part* of the spectrum. The part of the spectrum lying between will be called the *intermediate part* or the *transitional part.*

[2] Martin Hanke has also observed this phenomenon [personal communication].

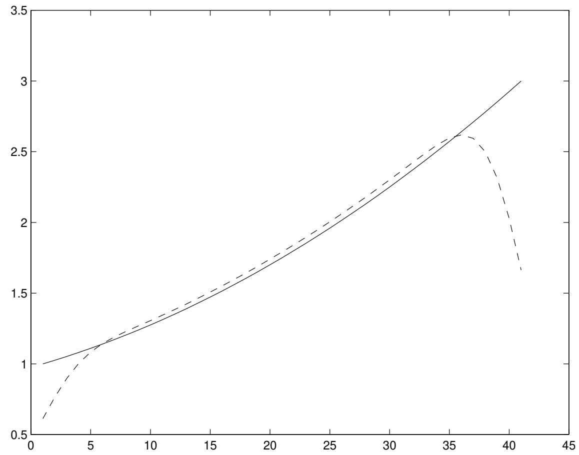

Figure 2.1: $x(t)$ (solid) and $b(t)$ (dashed)

2. An example

In this section we will introduce a simple-minded example from image processing. Let $\underline{x}(t)$ represent an "image" on $[0,1]$, and consider the Gaussian blurring operator

$$\underline{b}(t) = \int_0^1 \underline{x}(s) e^{\left(\frac{s-t}{\sigma}\right)^2} ds.$$

We will discretize this operator by choosing an integer n and setting

$$t_i = \frac{i-1}{n}, \qquad i = 1, \ldots, n$$

and generating a matrix K whose elements are

$$\kappa_{ij} = e^{\left(\frac{t_j - t_i}{\sigma}\right)^2}.$$

The resulting matrix is then scaled so that its dominant eigenvalue is one.

In the following experiments, we took $n = 40$, and $\sigma = 0.1$. Our function was $\underline{x}(t) = 1 + t + t^2$, and it was discretized by evaluating it at the points t_i to give the components x_i of x. The vector b was generated in the form Ax. Figure 2.1 is a plot of $\underline{x}(t)$ and $\underline{b}(t)$. The divergence at the ends of the intervals is due to the fact that at those points the part of the blurring distribution that lies outside $[0,1]$ becomes significant.

Figure 2.2 shows the vectors b, $\tilde{b}$, and x after they have been transformed to the spectral coordinates. Note that for $i > 15$ values $\tilde{\beta}_i$ are significantly contaminated with

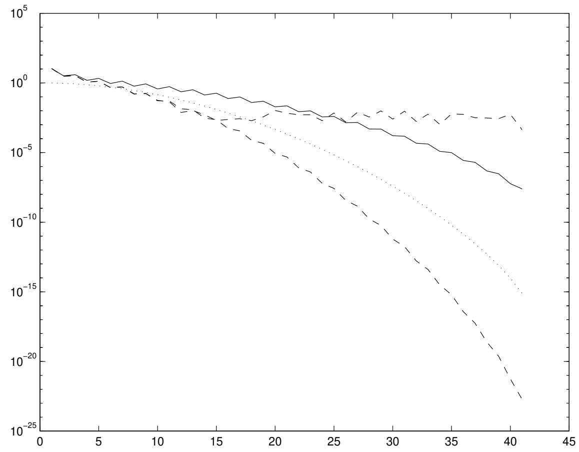

Figure 2.2: ξ_i (solid), λ_i (dotted) and β_i, $\tilde{\beta}_i$ (dashed)

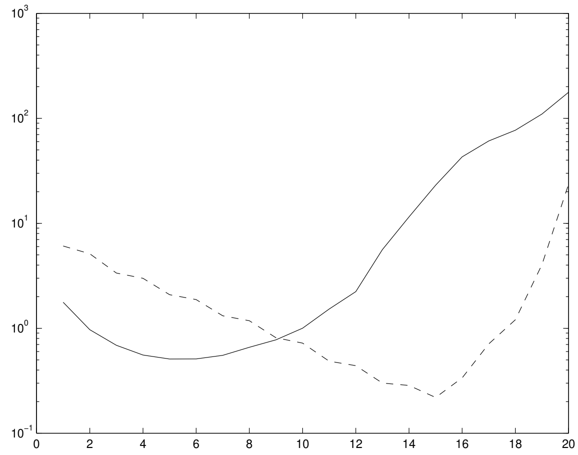

Figure 2.3: Errors from spectral regularization (dashed) and MINRES regularization (solid)

error. The plot also shows the eigenvalues of A, which decrease slowly at first and then rapidly plunge to about 10^{-15}. This rapid decrease is typical of smooth operators.

Truncated spectral regularization and MINRES regularization behave quite differently for this problem. Figure 2.3 plots the norms of the errors at the kth step of each method. The error for truncated spectral regularization decreases slowly until $k = 15$,

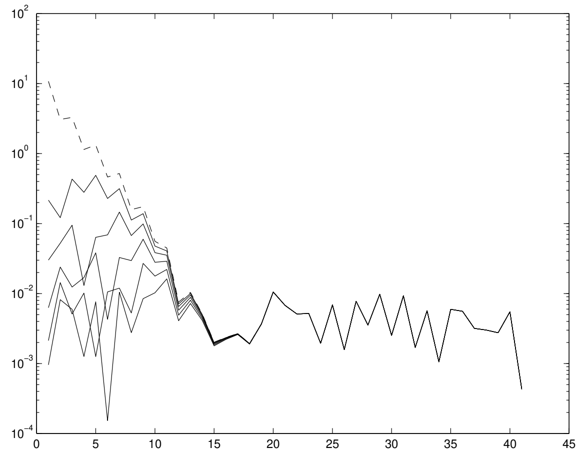

Figure 2.4: First five MINRES residuals and $\tilde{b}$ (dashed line)

after which it increases rapidly. The error in the MINRES iterates decreases rapidly until $k = 5$ and then rises. Truncated spectral regularization produces a better solution with about half the error, but arrives at it more slowly.[3] In fact, not until $k = 11$ does it produce a solution as accurate as the MINRES solution for $k = 5$.

It is informative to examine the behavior of the MINRES residuals. Figure 2.4 displays the magnitudes of the residuals at the eigenvalues λ_i as a function of i for the first five MINRES approximations along with the magnitudes of the components of $\tilde{b}$. It is seen that they decrease at the initial points of the spectrum. At other points in the spectrum, they change very little. The plot in Figure 2.5 shows the same thing in terms of the residual polynomials that satisfy (1.2).[4] The polynomials are small in the initial part of the spectrum and then rise swiftly to a value of one.

Stepping back from this example, we see that three factors are operating to give the fast regularization.

1. From (1.1) it follows that the residual cannot be reduced indefinitely without harming the solution.

[3]The relative quality of the solutions depends on the throw of the dice in constructing the random vector e. For ten other simulations the ratio of the minimum error for truncated spectral regularization to that for MINRES was the following: 0.80, 0.57, 0.76, 0.41, 1.14, 0.65, 0.67, 0.67, 0.82, 0.68. These numbers show the solutions to be usually comparable, with the MINRES solution actually better in a single case.

[4]The plots are not strictly the polynomials. The values between i and $i+1$ were obtained by evaluating the polynomial at ten equally spaced points between λ_i and λ_{i+1}.

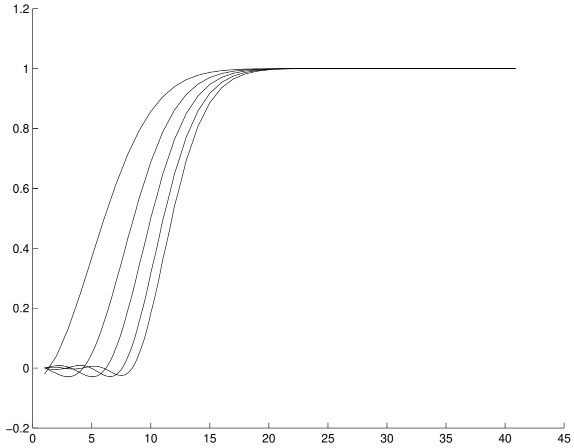

Figure 2.5: First five MINRES residual polynomials

2. The plot in Figure 2.4 shows that the residual decreases rapidly to a point where further reduction is harmful.

3. Notice in Figure 2.4 that the values of the first five residuals corresponding to terminal eigenvalues λ_i coincide with $\tilde{\beta}_i$. Equation (1.1) implies that if we reduce the residual at λ_i from $\tilde{\beta}_i$ to $\tilde{\beta}_i(1-\alpha)$, where $0 < \alpha < 1$, then the corresponding component of the approximate solution is $\alpha\tilde{\beta}_i/\lambda_i$. Now in the terminal part of the spectrum $\tilde{\beta}_i \cong \epsilon_i$ and hence the component of the approximate solution is essentially $\alpha\epsilon_i/\lambda_i$. Because the terminal eigenvalues are extremely small compared to the error, even a very small reduction of the residual at a terminal eigenvalue must result in a large component in the approximate solution. If there is any hope of MINRES producing an acceptable regularized solution, this must not occur within the first few iterations. As it turns out (see Figure 2.5), the first 5 residual polynomials for our running example are very near one on the terminal spectrum, so that there is virtually no reduction of the residual there.

In the next three sections we will establish results that give quantitative substance to these observations.

3. The residual and the error

The first of our results concerns the relation between the size of the error in a purported solution and the size of the corresponding residual. We have already noted that if the residual is required to be too small, the error must be large. However, we can be more precise.

Specifically, let y be a purported solution with residual norm δ. There are many such solutions, and among them one must give minimum error; i.e., there must be a vector y that solves the following problem. Given $\delta > 0$

$$\begin{array}{l} \text{minimize} \|x - y\|^2 \\ \text{subject to} \|\tilde{\beta} - \Lambda y\|^2 = \delta^2. \end{array} \tag{3.1}$$

The following theorem shows when the solution to this problem has a large error. Below we have assumed $\lambda_1 = 1$.

Theorem 3.1. *If $\tau > -1$ is the solution of the equation*

$$\sum_i \frac{\epsilon_i^2}{(1 + \tau\lambda_i^2)^2} = \delta^2, \tag{3.2}$$

then

$$\eta_i = \xi_i + \frac{\tau\lambda_i\epsilon_i}{1 + \tau\lambda_i^2} \tag{3.3}$$

is a solution of (3.1). *In this case*

$$\|x - y\|^2 = \sum_i \left(\frac{\tau\lambda_i}{1 + \tau\lambda_i^2}\right)^2 \epsilon_i^2. \tag{3.4}$$

Proof. Consider the Lagrangian

$$\sum_i (\xi_i - \eta_i)^2 + \tau \left[\sum_i (\tilde{\beta}_i - \lambda_i\eta_i)^2 - \delta^2\right],$$

where τ is the Lagrange multiplier. Differentiating with respect to η_i and setting the results to zero we get

$$\xi_i - \eta_i + \tau\lambda_i(\tilde{\beta}_i - \lambda_i\eta_i) = 0.$$

Equation (3.3) follows on solving this equation for η_i and using the fact that $\tilde{\beta}_i = \lambda_i\xi_i + \epsilon_i$. Equation (3.2) defines the value of τ for which the solution satisfies the constraint. Finally (3.4) follows by direct computation. ∎

The best way to understand this theorem is to consider the solution y_τ generated as τ varies from -1 to ∞. From (3.2) we see that the residual norm δ decreases monotonically from ∞ to 0. The error norm $\|x - y_\tau\|$ decreases from ∞ to 0 at $\tau = 0$ where $\delta = \|e\|$. Thereafter small reductions in the residual cause large increases in the error. To see this, consider the term

$$\frac{\epsilon_i^2}{(1+\tau\lambda_i^2)^2}$$

in (3.2). For this term to be reduced by a factor of four from its value ϵ_i^2 at $\tau = 0$, a very modest reduction in the total, we must have $\tau\lambda_i^2 = 1$. The corresponding term in the square of the error norm (3.4) is

$$\left(\frac{\tau\lambda_i\epsilon_i}{1+\tau\lambda_i^2}\right)^2 = \frac{\epsilon_i^2}{4\lambda_i^2},$$

which for small λ_i is large. It is worth noting that these observations depend only on Λ and e, not on x or b.

The above comments suggest that an attempt to reduce the residual norm below $\|e\|$ will increase the error in the solution. This is certainly true of our example, for which $\|e\| = 0.033$. The error turns around at $k = 5$. The residual norms are $\rho_4 = 0.055$ $\rho_5 = 0.039$, and $\rho_6 = 0.031$. Thereafter the error increases sharply, as it must (see Figure 2.3).

It is worth noting that Theorem 3.1 supports the discrepancy principle of Morozov, which says that a regularization parameter should be chosen to make the norm of the residual approximately the size of the norm of the error (see [4, §4.3]). It also suggests that it is better to make the residual too large rather than too small.

4. Reduction of the residual

We now turn to the reduction of the residual. The usual analysis of MINRES is based on (1.2). Specifically, given any polynomial p_k of degree k satisfying $p_k(0) = 1$, we have

$$\|r_k\|^2 \le \sum_{i=1}^{n} p_k(\lambda_i)^2\tilde{\beta}_i^2. \tag{4.1}$$

One convergence result amounts to choosing a sequence of polynomials p_k such that

$$\lambda \in [\lambda_n, \lambda_1] \implies p_k(\lambda) \to 0.$$

The rate at which the p_k approach zero on $[\lambda_n, \lambda_1]$ is an upper bound on the convergence rate of MINRES.

The key to analyzing the regularizing properties of MINRES is to note that we are not interested in the convergence of the method on the terminal part of the spectrum. In fact

(1.1) shows that MINRES *must not* converge there until after a reasonable regularized solution has been formed. We can therefore restrict our attention to an interval, say, $[\lambda_m, \lambda_1]$ in which the $\tilde{\beta}_i$ are relatively error free. Since the interval is farther removed from the origin, the convergence will be faster.[5]

The polynomial $p_{k,m}$ for this analysis is constructed in the usual way from the Chebychev polynomials $c_k(\tau)$. Specifically, let

$$s_m(t) = -\frac{2t - \lambda_1 - \lambda_m}{\lambda_1 - \lambda_m} \tag{4.2}$$

and set

$$p_{k,m}(t) = \frac{c_k[s_m(t)]}{c_k[s_m(0)]}, \tag{4.3}$$

By bounding the values of these polynomials we obtain the following theorem.

Theorem 4.1. *Let* $m > 1$ *and set*

$$\kappa = \frac{\lambda_1}{\lambda_m}.$$

and

$$\tilde{b}_m = (\tilde{\beta}_1 \ \ldots \ \tilde{\beta}_m)^{\mathrm{T}}$$

Then

$$\begin{aligned} \|r_k\|^2 &\leq \|p_{k,m}(\Lambda)\tilde{b}\|^2 \\ &\leq 4\left(\tfrac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k} \|\tilde{b}_m\|^2 \\ &\quad + \textstyle\sum_{i=m+1}^{n} \left(1 - 2\tfrac{\lambda_i}{(\sqrt{\lambda_1}+\sqrt{\lambda_i})^2}\right)^{2k} \tilde{\beta}_i^2. \end{aligned} \tag{4.4}$$

The proof of this theorem is highly technical and will be found in appendix 9. The first term in (4.4) is of the same form as the bound from the usual analysis of MINRES mentioned above. However, for that bound κ is λ_1/λ_n, which is larger than λ_1/λ_m and implies slower convergence. The first factor in the remaining terms is essentially one for small eigenvalues. However, for eigenvalues near λ_m it has a reducing effect on the corresponding part of the residual.

[5] For the standard analysis see [5, Ch. 3]. It should not be thought that the idea of varying the polynomial is new: e.g., see [5, pp. 53–54] in which a special polynomial is used to take advantage of a well separated dominant eigenvalue. What is new in this paper is the form of the polynomial, which is adapted to the analysis of iterative regularization.

The bound is too crude to use in applications, but it provides insight into the way MINRES converges. In our running example, when m is small, κ is near one, and the polynomials $p_{k,m}$ rapidly reduce the residual on $\lambda_1, \ldots, \lambda_m$. But this is precisely the place where there is a large residual to reduce. Since the reduction by $p_{k,m}$ bounds the MINRES reduction, we conclude that initially MINRES will rapidly reduce the residual.

But the bound implies more. As k increases, for fixed m, $p_{k,m}$ will reach a point of diminishing returns when there is not enough residual to reduce. This can be seen by looking at the upper bound in (4.4). Consider m fixed at 4. Since κ is effectively 1, the first term becomes insignificant for $k \geq 2$ and it turns out we have a tight bound on the residual for $k = 2$. Yet as k increases and the residual continues to be reduced, the bound stagnates because the second term decreases slowly. However, if we increase m a little the first term remains small while the second term becomes smaller (since there are fewer indices over which to sum). Thus with increasing k one should increase m to get the smallest possible upper bound. The following table shows that this is precisely what happens in our example.

k	1	2	3	4	5
m	2	4	5	6	7
$\|p_{k,m}(\Lambda)\tilde{b}\|$	8.8e−01	2.8e−01	1.3e−01	7.1e−02	4.7e−02
$\max_{[\lambda_m,\lambda_1]} p_{k,m}$	8.0e−02	2.9e−02	1.0e−02	6.1e−03	5.2e−03
ρ_k	8.7e−01	2.5e−01	1.0e−01	5.5e−02	3.9e−02

For each value of k, the value of m for which $p_{k,m}$ minimizes the residual $\|p_{k,m}(\Lambda)\tilde{b}\|$ was determined. As the table shows, the value of m increases with k. This increase is accompanied by a modest decrease in the size of $p_{k,m}$ on $[\lambda_m, \lambda_1]$. Note that by holding m fixed one can obtain a dramatic decrease in the size of $p_{k,m}$ over $[\lambda_m, \lambda_1]$, but this strategy is not as effective at reducing the bound on the residual norm as the strategy of increasing m and putting up with a modest decrease in $p_{k,m}$. For example, if m is fixed at two, then $\max_{[\lambda_2,\lambda_1]} p_{5,2} = 1.5\text{e}{-09}$, but the residual norm is only 8.9e−02.

The last line in the table shows the MINRES residual norms. They are well approximated by the values in the third line, which suggests that MINRES is following a similar strategy. Figure 2.5 supports our contention that MINRES reduces the residual by increasing the interval in which the polynomial is small.

5. Nonreduction of the residual

Let $\alpha \in [0, 1]$, and suppose we wish to reduce the residual at λ_i from $\tilde{\beta}_i$ to $\tilde{\beta}_i(1 - \alpha)$. We have already noted that the corresponding component of $\tilde{x}$ is then

$$\tilde{\xi}_i = \frac{\alpha \tilde{\beta}_i}{\lambda_i}.$$

If λ_i is small, then $\tilde{\xi}_i$ is large and the solution will be unsatisfactory. Thus while MINRES is producing a regularized solution, it must not reduce the residuals on the terminal part of the spectrum. From Figure 2.5 we see that this is precisely what is happening in our example. The residual polynomials are effectively one at small λ_i.

The reason MINRES behaves in this manner is that if any polynomial is small at a small eigenvalue, it must be very large at the initial eigenvalues, say $\lambda_1, \ldots, \lambda_m$. Specifically, we have the following theorem.

Theorem 5.1. *Let $p_{k,m}$ be defined by (4.3). If p is any polynomial of degree k satisfying $p(0) = 1$ and $p(\lambda_i) = 1 - \alpha$ for some $(i > m)$, then there is a point $t_{\text{big}} \in [\lambda_m, \lambda_1]$ such that*

$$|p(t_{\text{big}})| \geq \left| 1 - \frac{p_{k-1,m}(\lambda_i) - 1 + \alpha}{p_{k-1,m}(\lambda_i)} \frac{\lambda_m}{\lambda_i} \right| |p_{k-1,m}(\lambda_m)|. \tag{5.1}$$

Proof. The proof is a variant of the standard proof of the theorem of de laVallée Poussin [1, p. 191]. Consider the polynomial

$$q(t) = \left[1 - \frac{p_{k-1,m}(\lambda_i - 1 + \alpha)}{p_{k-1,m}(\lambda_i)} \frac{t}{\lambda_i} \right] p_{k-1,m}(t).$$

This polynomial satisfies the conditions of the theorem. Moreover, it alternates in sign at k points

$$\lambda_m = t_1, t_2, \ldots, t_k = 1$$

on the interval $[\lambda_m, \lambda_1]$.

Now let p also satisfy the conditions of the theorem. We claim that $|p(t_i)| \geq |q(t_i)|$ for at least one t_i. For if not, the polynomial $p - q$ alternates in sign at the t_i and hence has $k - 1$ zeros in the interval $[\lambda_m, \lambda_1]$. But by the conditions of the theorem $p - q$ is zero at zero and λ_i. Hence $p - q$ is a polynomial of degree not greater than k with $k + 1$ zeros and must be identically zero — a contradiction.

The bound (5.1) now follows on evaluating q at λ_m, where the value of $|q(t_i)|$ is the smallest. ∎

The import of this theorem is that the cost of reducing the residual at the small eigenvalues is to make the polynomial large on an interval containing the initial eigenvalues. This is because the factor λ_m/λ_i in (5.1) becomes enormous for i large. For example, Theorem 5.1 implies that if $m = 8$ and $k = 5$ in our running example then to get a reduction of .9 ($\alpha = .1$) in the residual at $\lambda_{25} = 6.8\text{e}{-6}$, the absolute value of the residual polynomial will have to be to be greater than 71 at some point in $[\lambda_8, \lambda_1]$. Such a large polynomial is more likely to cause a residual magnification than a residual reduction in $[\lambda_8, \lambda_1]$. This likelihood is increased by the fact that for our example the

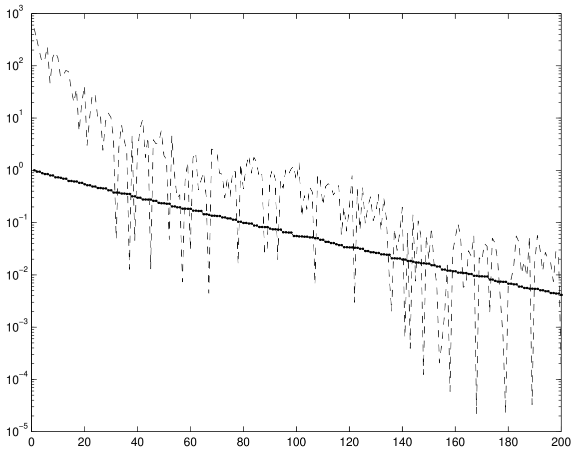

Figure 6.1: Eigenvalues (solid line) and components of b (dashed line)

roots of p_k are spread out over $[\lambda_m, \lambda_1]$, where $m > k$, and the zeros of p_k cannot be near all m of these eigenvalues.

On the other hand, in our example suppose we wish to reduce the residual at λ_{15} from $\tilde{\beta}_{15}$ to $\tilde{\beta}_{15}(1-\alpha)$ for $\alpha = 0.11$. The lower bound (5.1) for the residual polynomial is 0.0086. The maximum value of the residual polynomial in $[\lambda_8, \lambda_1]$ is, in fact, 0.037, which is insufficiently greater than 0.0086 to render the reduction at λ_{15} harmless. Once again, the dispersal of the roots of $p_{k,m}$ has a beneficial effect — this time allowing a modest decrease in the residual at the intermediate eigenvalues.

In the last section we observed that MINRES eschews a quick reduction of the residual polynomials over a limited range in favor of expanding the range in which they are moderately small. The theory of this section shows that the deferred reduction has another benefit: it gives the algorithm an opportunity to reduce the residual by a modest amount on the intermediate eigenvalues.

6. Another example

To confirm our analysis of MINRES, we consider a two dimensional blurring operator. Specifically, the original matrix A is the Kronecker product $T \otimes T$, where T is a symmetric Toeplitz matrix whose first row is given by

$$t_{1k} = e^{-0.1(1-k)^2}, \qquad k = 1, \ldots, 32.$$

Figure 6.1 exhibits the eigenvalues of the matrix and the absolute values components of

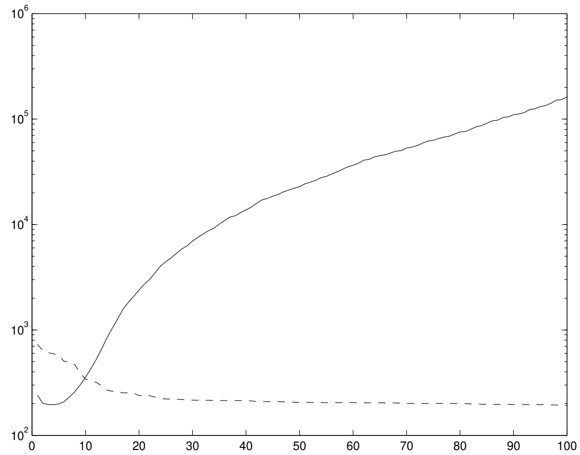

Figure 6.2: Errors from spectral regularization (dashed line) and MINRES regularization (solid line)

b in the spectral coordinate system. The error e consisted of white noise scaled so that $\|e\|/\|b\| = 0.01$.

Figure 6.2 plots the errors from spectral regularization and MINRES. The contrast is dramatic. The MINRES iteration reduces the error to a minimum of 195 in four iterations. Spectral regularization requires the first 100 components of the solution be computed to reduce the error to a minimum of 192. (However, the plot shows that acceptable accuracy is attained for, say, 30 components.)

The analysis of Section 3 suggests that the solution will deteriorate when the residual norm becomes less than the error norm. In fact the fourth, fifth, and sixth residual norms are 8.66, 7.86, and 7.52. This should be compared with the error norm 7.78.

From the comments in Section 4 we would expect that residual polynomials would reduce the error by spreading their roots over a number of eigenvalues rather than by approximating specific eigenvalues. The roots for the first five residual polynomials are given below.

1	9.3e−01				
2	9.8e−01	7.3e−01			
3	9.9e−01	8.1e−01	5.7e−01		
4	9.9e−01	8.7e−01	6.9e−01	3.6e−01	
5	1.0e+00	9.0e−01	7.4e−01	5.4e−01	2.2e−01

The largest root is converging to one, which in our normalization is an eigenvalue of A. But otherwise, the roots are dispersed. For example, the smallest root of p_4 is 0.36, which corresponds to λ_{34}.

7. Conclusions

In the introduction to this paper we emphasized that this work would not provide a complete analysis of MINRES as a regularization technique. We can analyze MINRES applied to well posed problems because asymptotic bounds become applicable long before a solution of the desired accuracy is attained. On the other hand, for ill-posed problems in our examples MINRES does not have a chance to settle down before the optimal solution is obtained. In effect, we are analyzing the early transient behavior of the algorithm — always a difficult problem.

In spite of not concluding with a general theorem, our analysis sheds considerable light on the behavior of MINRES as a regularizer. Theorem 3.1 is a very general result about ill-posed problems. It suggests that attempting to reduce the residual norm below the error norm will result in a worsened solution. As we have already noted the theorem can be taken as a justification of the Morozov discrepancy principle.

We have mentioned that the bound (4.4) on the residual norm is too crude to be used in practical applications. Nonetheless, it is good enough to suggest the strategy that MINRES uses to obtain rapid reduction of the residual norm. The free parameter m in the bound is essentially a measure of the spread of the zeros of the residual polynomial. As we have seen, the bound is optimized by keeping m greater than the iteration number k. This suggests that MINRES will attempt to minimize the residual by spreading the roots of its residual polynomials over a larger number of eigenvalues. Our numerical examples support this conjecture.

Finally, Theorem 5.1 shows us why MINRES does not tend to reduce the residual at a small eigenvalue while homing in on the regularized solution. Essentially, a small value of the residual polynomial at a small eigenvalue causes the polynomial to be large on the larger eigenvalues — just where it is most desirable to obtain a decrease in the residuals.

For definiteness, we have restricted ourselves to MINRES applied to a positive definite system. Obviously, some of our results apply to other systems and methods — e.g., least squares regularization of overdetermined systems. Such generalizations, however, are the subject for future work.

8. Acknowledgements

We with to thank Per-Christian Hansen and Dianne O'Leary for their useful comments on the first draft of this paper.

9. Appendix

Proof of Theorem 4.1: Let $c_k(\tau)$ denote the Chebychev polynomial of degree k and define $s_m(t)$ and $p_{k,m}(t)$ as in (4.2) and (4.3), respectively. Now $p_{k,m}$ is a polynomial of degree k satisfying $p_{k,m}(0) = 1$, and therefore from (4.1), we have

$$\|r^{(k)}\|_2^2 \le \sum_{i=1}^{n} p_{k,m}^2(\lambda_i)\tilde{\beta}_i = \|p_{k,m}(\Lambda)\tilde{b}\|_2^2. \tag{9.1}$$

When $i \le m$, $c_k^2(s_m(\lambda_i)) \le 1$. Moreover, one can show that the denominator of $p_{k,m}$, which is independent of i, is bounded above [9]:

$$\left| \frac{1}{c_k(\frac{\lambda_1+\lambda_m}{\lambda_1-\lambda_m})} \right| \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k. \tag{9.2}$$

Thus, it remains to determine an upper bound for the values $c_k(s_m(\lambda_i)),\ i > m$.

For notational convenience, we fix i and let $d = \frac{\lambda_i}{\lambda_1}$, and note that since $\kappa = \frac{\lambda_1}{\lambda_m}$, $\frac{\lambda_i}{\lambda_m} = d\kappa$. Using this notation, we write $s_{m,i} \equiv s_m(\lambda_i) = \left(\frac{\kappa+1-2d\kappa}{\kappa-1}\right)$.

Now $c_k(s) = \cosh(k\cosh^{-1} s)$. We use the formulas

$$\cosh x = \frac{e^x + e^{-x}}{2} \tag{9.3}$$

and

$$\ln z = \cosh^{-1}\left(\frac{z+z^{-1}}{2}\right) \text{ when } x = \ln z \tag{9.4}$$

to determine our upper bound as outlined below. We note that this type of proof technique was used at least as early as the 1960's (see [3]).

1. Find a z such that $\left(\frac{z+z^{-1}}{2}\right) = s_{m,i}, i > m$.
2. Use z and (9.3) to determine $c_k(s_{m,i}), i > m$.
3. Find an upper bound for $c_k^2(s_{m,i}), i > m$.
4. Use this bound and (9.2) to deduce an upper bound for $p_{k,m}^2(s_{m,i})\ i > m$.
5. Use the above information to obtain the upper bound on the right hand side of (9.1).

Step 1: We first set

$$\frac{z+z^{-1}}{2} = s_{m,i}$$

to obtain

$$z^2 - 2s_{m,i}z + 1 = 0.$$

Since this is a quadratic equation, z is given by

$$\begin{aligned} z &= \frac{2s_{m,i} \pm \sqrt{4s_{m,i}^2 - 4}}{2} \\ &= \left(\frac{\kappa+1-2d\kappa}{\kappa-1}\right) \pm \sqrt{\left(\frac{\kappa+1-2d\kappa}{\kappa-1}\right)^2 - 1} \\ &= \frac{1}{\kappa-1}\left(\kappa+1-2d\kappa \pm 2\sqrt{(\kappa-\kappa d)(1-\kappa d)}\right) \\ &= \frac{(\sqrt{\kappa-\kappa d} \pm \sqrt{1-\kappa d})^2}{\kappa-1}. \end{aligned} \tag{9.5}$$

For convenience, we take z to be

$$z = \frac{(\sqrt{\kappa-\kappa d} + \sqrt{1-\kappa d})^2}{\kappa-1}. \tag{9.6}$$

Step 2: Now $c_k(s_{m,i}) = \cosh(k \ln z) = \cosh(\ln z^k)$. Using (9.3), we obtain

$$\begin{aligned} c_k(s_{m,i}) &= \frac{z^k + z^{-k}}{2} \\ &= \frac{z^{2k}+1}{2z^k}. \end{aligned} \tag{9.7}$$

Step 3: Therefore,

$$c_k^2(s_{m,i}) = \frac{1}{4}\left(z^{2k} + \frac{1}{z^{2k}} + 2\right). \tag{9.8}$$

To get an upper bound on $c_k^2(s_{m,i})$, let us first try to get an upper bound on $\frac{1}{z^{2k}}$ by getting a lower bound on z. Observe that the numerator of z in (9.6) is bounded below by $\kappa - d\kappa$, so our lower bound for z is $\frac{\kappa-d\kappa}{\kappa-1}$. Since $d\kappa \le 1$, we therefore obtain

$$\frac{1}{z} \le \frac{\kappa-1}{\kappa-d\kappa} \le 1$$

and so $\frac{1}{z^{2k}} \leq 1$.

Next, we have the following expression for z^{2k}:

$$z^{2k} = \left(\frac{\kappa + 1 + 2\left((\kappa - d\kappa)(1 - d\kappa)\right)^{1/2} - 2d\kappa}{\kappa - 1} \right)^{2k}.$$

Therefore, we have the following bound on $c_k^2(s_{m,i})$:

$$c_k^2(s_{m,i}) \leq \frac{1}{4} \left(3 + \left(\frac{\kappa + 1 + 2\left((\kappa - d\kappa)(1 - d\kappa)\right)^{1/2} - 2d\kappa}{\kappa - 1} \right)^{2k} \right). \tag{9.9}$$

Step 4: Now we may use equations (9.2) and (9.9) to determine an upper bound on $p_{k,m}^2(s_{m,i})$ for $i > m$:

$$p_{k,m}^2(s_{m,i}) \leq \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} \left(3 + \left(\frac{\kappa + 1 + 2\left((\kappa - d\kappa)(1 - d\kappa)\right)^{1/2} - 2d\kappa}{\kappa - 1} \right)^{2k} \right)$$

Using $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} = \frac{\kappa - 1}{\kappa + 2\sqrt{\kappa} + 1}$, we cancel like terms in the right hand side above and factor out $\sqrt{\kappa}$ from the radical to obtain

$$p_{k,m}^2(s_{m,i}) \leq 3 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} + \left(\frac{\kappa + 1 + 2\sqrt{\kappa}\left((1 - \frac{\lambda_i}{\lambda_1})(1 - \frac{\lambda_i}{\lambda_m})\right)^{1/2} - 2\frac{\lambda_i}{\lambda_m}}{\kappa + 1 + 2\sqrt{\kappa}} \right)^{2k}.$$

The second term of the right hand side above can be bounded above by

$$\left(1 - \frac{2\frac{\lambda_i}{\lambda_m}}{\kappa + 1 + 2\sqrt{\kappa}} \right)^{2k} = \left(1 - 2\frac{\lambda_i}{\lambda_1 + \lambda_m + 2\sqrt{\lambda_1 \lambda_m}} \right)^{2k}$$

Thus,

$$p_{k,m}^2(s_{m,i}) \leq 3 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} + \left(1 - 2\frac{\lambda_i}{\lambda_1 + \lambda_m + 2\sqrt{\lambda_1 \lambda_m}} \right)^{2k}. \tag{9.10}$$

Step 5: From equation (9.2) when $i \leq m$, $p_{k,m}^2(s_{m,i}) \leq \left(\frac{l}{c_k(s_{m,i})} \right)^2 \leq 4 \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \right)^{2k}$. Breaking up the sum over the first m and last $n - m$ terms, from (9.1) and (9.10) we observe

$$\|r_k\|_2^2 \leq 4 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^{2k} \|\tilde{b}_m\|_2^2 + \sum_{i=m+1}^{n} \left(1 - 2\frac{\lambda_i}{(\sqrt{\lambda_1} + \sqrt{\lambda_m})^2} \right)^{2k} \tilde{\beta}_i^{\,2}.$$

References

[1] K. E. Atkinson. *An Introduction to Numerical Analysis.* John Wiley, New York, 1978.

[2] Å. Björck. *Numerical Methods for Least Squares Problems.* SIAM, Philadelphia, 1996.

[3] J. W. Daniel. The conjugate gradient method for linear and nonlinear operator equations. *SIAM Journal on Numerical Analysis*, 4:10–26, 1967.

[4] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[5] A. Greenbaum. *Iterative Methods for Solving Linear Systems.* SIAM, Philadelphia, 1997.

[6] P. C. Hansen. The truncated SVD as a method for regularization. *BIT*, 27:534–553, 1987. Cited in [2].

[7] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems.* SIAM, Philadelphia, 1998.

[8] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis*, 12:617–629, 1975.

[9] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer-Verlag, New York, second edition, 1993.