

PITFALLS IN THE NUMERICAL SOLUTION
OF LINEAR ILL-POSED PROBLEMS

J.M. Varah
Computer Science Department
University of British Columbia

TECHNICAL REPORT 81-10

ABSTRACT

Very special computational difficulties arise when attempting to solve linear systems arising from integral equations of the first kind. We examine here existence and uniqueness questions associated with so-called reasonable solutions for such problems, and present results using the best-known methods on inverse Laplace transform problems. We also discuss the choice of free parameters occurring in these methods, from the same point of view.

PITFALLS IN THE NUMERICAL SOLUTION
OF LINEAR ILL-POSED PROBLEMS

J.M. Varah
Computer Science Department
University of British Columbia

TECHNICAL REPORT 81-10

ABSTRACT

Very special computational difficulties arise when attempting to solve linear systems arising from integral equations of the first kind. We examine here existence and uniqueness questions associated with so-called reasonable solutions for such problems, and present results using the best-known methods on inverse Laplace transform problems. We also discuss the choice of free parameters occurring in these methods, from the same point of view.

1. INTRODUCTION

We are concerned with the numerical solution of the linear $n \times n$ system

$$Kf = g \quad (1.1)$$

where K is inherently ill-conditioned because of the source of the system. We assume for definiteness that the source is an integral equation of the first kind in one dimension:

$$\int \tilde{K}(s,t) \tilde{f}(t) dt = \tilde{g}(s). \quad (1.2)$$

We shall only consider the case when $\tilde{K}$ is a smooth kernel; then clearly the mapping $\tilde{f} \rightarrow \tilde{g}$ by (1.2) "smooths out" functions, and even takes non-continuous $\tilde{f}$ into smooth $\tilde{g}$, so we can't hope to solve (1.2) for arbitrary $\tilde{g}$. This also applies to (1.1), of course, and in Section 2 we examine this problem in more detail, in an attempt to specify when the given discrete problem (1.1) has a reasonable solution.

Several methods have been developed for solving (1.1), and in Section 3 we consider four of these: the truncated singular value decomposition method, the regularization method, the modified regularization method, and a function expansion method. We refer to Varah (1979) and Bjorck and Elden (1979) for more details and references. Each of these methods involves a free parameter, and for appropriate choice of this parameter, each method is relatively stable with respect to perturbations in the data, so that each is a reasonable

computational method. As we shall see, however, the solutions obtained may be very different from one another, so that it is extremely difficult to say which is "the best" numerical solution. In Section 4, we examine several numerical examples illustrating this, all involving the inverse Laplace transform operator in (1.2).

Finally in Section 5, we discuss the choice of free parameters in the methods. We find that although reasonable choices can be made for the expansion methods, this does not appear to be the case for regularization methods.

2. EXISTENCE OF REASONABLE SOLUTIONS

For the continuous problem (1.2), one way of specifying the existence of a reasonable solution is the Picard condition: for L_2 kernels $\tilde{K}$ with $\iint \tilde{K}(s,t)^2 ds dt < \infty$, there are orthogonal functions $\{\phi_i(s)\}$, $\{\psi_i(t)\}$, and corresponding scalars $\lambda_i \rightarrow 0$ so that

$$\int \tilde{K}(s,t) \phi_i(s) ds = \lambda_i \psi_i(t)$$

$$\int \tilde{K}(s,t) \psi_i(t) dt = \lambda_i \phi_i(s).$$

Then if $\tilde{g}(s) = \sum \tilde{\beta}_i \phi_i(s)$, then $\tilde{f}(t) = \sum (\tilde{\beta}_i / \lambda_i) \psi_i(t)$; however $\tilde{f} \in L_2$ only if $\sum (\tilde{\beta}_i / \lambda_i)^2 < \infty$, which is the Picard condition. This obviously restricts the class of data function $\tilde{g}$ for the problem.

We should add at this point that the rate at which the $\lambda_i \rightarrow 0$ depends directly on the smoothness of the operator $\tilde{K}$: for example, when $\tilde{K}$ is the (non-smooth) Green's function

$$\tilde{K}(s,t) = \begin{cases} s(1-t), & s < t \\ t(1-s), & s \geq t, \end{cases}$$

then $\lambda_n = \frac{\pi^2}{n^2}$. And, when $\tilde{K}$ is the (smooth) harmonic continuation operator

$$\tilde{K}(s,t) = \frac{1}{2\pi} \frac{1-\rho^2}{1-2\rho \cos(s-t)t_\rho^2}$$

on $(0, 2\pi)$, then $\lambda_n = \rho^n$ ($\rho < 1$). Thus, smooth kernels lead to much more ill-conditioned problems.

Now consider the discrete problem (1.1), which can be derived from (1.2) by applying a specific quadrature rule in t , say

$$\sum_{j=1}^n w_j \tilde{K}(s, t_j) \tilde{f}(t_j) = \tilde{g}(s)$$

which, applied at n sample points s_i gives the linear system $Kf = g$. The discrete analogue of the above expansions for the continuous problem is of course the singular value decomposition (SVD)

$$K = UDV^T,$$

where U and V are orthogonal, and $D = \text{diag}(\sigma_i)$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$. However, to derive some reasonable discrete version of the Picard condition is not so straightforward - we need first of all to define what is meant by a reasonable solution to (1.1).

Definition: The vector f is a reasonable solution to (1.1) with noise level ϵ if $\|f\| = O(1)$ and $\|Kf - g\| = O(\epsilon)$.

Here we should mention that the $O(\epsilon)$ notation means "about the same size as" and should not be taken to mean anything in an asymptotic sense.

The question of whether a given discrete problem (1.1) has a reasonable solution can be decided by means of the SVD of K : if $K = UDV^T$ and we define y and β by $V^T f = y$, $U^T g = \beta$, then

$$||Kf-g||_2^2 = \sum_1^n (\sigma_i y_i - \beta_i)^2 \text{ and } ||f||_2^2 = \sum_1^n y_i^2.$$

Thus it is clear that for f to be a reasonable solution, the given data g must be such that

$$\beta_i = O(\sigma_i), \quad i \leq i_0, \quad (2.1)$$

with i_0 chosen so that $|\beta_i| < \epsilon$ for $i > i_0$.

Notice this is a condition on the discrete data g , and is in some sense a discrete analogue of the Picard condition. Typically in practice, if the data have noise level ϵ , the $|\beta_i|$ decrease to $O(\epsilon)$ and remain at about this level for $i_0 < i \leq n$, whereas the σ_i decrease to zero. Thus for there to be a reasonable solution, the $|\beta_i|$ must decrease as fast as the σ_i , down to the noise level.

Also, this condition (2.1) provides a basis for existence of a reasonable solution f , but says nothing about uniqueness. Indeed, it is clear again from the SVD expansion that there will be many reasonable solutions; for $i > i_0$ (assuming $\sigma_i < \epsilon$ for $i > i_0$ as well) the y_i are essentially arbitrary. We will return to this point later when comparing the solutions obtained by various methods. Here we wish to emphasize that all such solutions are equally valid from a strictly computational point of view; in order to obtain unique solutions, we must further restrict the problem using other means (for example by restricting the class of solutions allowed or specifying a particular expansion for the solution).

We also feel this approach (of reasonable solutions) is more appropriate than that of conditioning given in Varah (1979), because it does not depend

on the method used to solve the problem.

3. THE NUMERICAL METHODS

(a) SVD: this is the truncated SVD method, which produces f as $f = \sum_{i=1}^k y_i v^{(i)}$, where $y_i = \beta_i / \sigma_i$ and k is the free parameter. If k is chosen as is (as in Section 2) with noise level ϵ known, and if the problem has a reasonable solution, then it is easy to see that this always produces such a reasonable solution: $\|Kf - g\|_2 = O(\epsilon)$ and $\|f\|_2 = O(1)$. Since $y_i = 0$ for $i > i_0$, the SVD solution can be said to provide the minimal reasonable solution possible.

Moreover, even if the noise level ϵ is not known, one can examine the sequences $\{\sigma_i\}$, $\{\beta_i\}$ and take k , the cutoff point, where the σ_i become smaller than the $|\beta_i|$. This gives a reasonable solution for ϵ chosen accordingly.

(b) LS: this is the original regularization method, which produces f as

$$f_\alpha = \min_f (\|Kf - g\|_2^2 + \alpha^2 \|f\|_2^2)$$

or equivalently as the solution of the normal equations

$$(K^T K + \alpha^2 I) f_\alpha = K^T g.$$

In examining the nature of the solution f_α , it is most convenient to again use the SVD expansion; then

$$f_\alpha = \sum_{i=1}^n \left(\frac{\beta_i}{\sigma_i + \alpha^2 / \sigma_i} \right) v^{(i)}. \quad (3.1)$$

At first sight, it appears that this is rather different from the SVD solution in (a), since all vectors $v^{(i)}$ contribute. However for properly chosen values of the free parameter α , most of these are damped, and we are left with a solution very much like the SVD. To see this, consider the same situation as before, with $\sigma_i \rightarrow 0$, $|\beta_i| \rightarrow \epsilon$ and $\beta_i = 0(\epsilon)$ for $i > i_0$. First, if $\alpha \ll \epsilon$ and we have some $\sigma_i = \alpha$, then the i -th term in (3.1) is $\frac{\beta_i}{2\alpha} v^{(i)}$ and $\left| \frac{\beta_i}{2\alpha} \right| \gg 1$ so we cannot have a reasonable solution - so we must choose $\alpha \geq \epsilon$. Moreover if α is chosen roughly equal to ϵ , then for $i \leq i_0$, the i -th coefficient $\frac{\beta_i/\sigma_i}{1+\alpha^2/\sigma_i^2} \simeq \beta_i/\sigma_i$ since $\sigma_i \geq \epsilon$ and for $i > i_0$, $\frac{\beta_i/\sigma_i}{1+\alpha^2/\sigma_i^2} \simeq 0$ since $\sigma_i \rightarrow 0$.

Thus the LS solution, with appropriate α , is close to the SVD solution, and this is certainly borne out by numerical experience. The most important connection is that both give a solution in terms of the expansion of singular vectors $\{v^{(i)}\}$ of K .

(c) MLS: this is the modified regularization method, which produces f as

$$f_\alpha = \min_f (||Kf-g||_2^2 + \alpha^2 ||Lf||_2^2)$$

or equivalently as the solution of the normal equations

$$(K^T K + \alpha^2 L^T L) f_\alpha = K^T g.$$

Here L is usually some discrete approximation to a derivative operator, and we in fact use simple forward differences to get L .

Here the appropriate expansion is not the SVD of K , but the generalized SVD of the matrix pair (K, L) as in van Loan (1976):

$$K = UD_aX^{-1}, \quad L = VD_bX^{-1},$$

with U and V orthogonal, D_a and D_b diagonal, and X the matrix which simultaneously diagonalizes the symmetric pair (K^TK, L^TL) . Using this decomposition, the solution f_α can be written as an expansion in the vectors $x^{(i)}$, the columns of X :

$$f_\alpha = \sum_1^n \left(\frac{\beta_i}{a_i + \alpha^2 b_i^2 / a_i} \right) x^{(i)}$$

where again $\beta = U^T g$ and $\{a_i\}$, $\{b_i\}$ are the elements of D_a , D_b , ordered so that the $\{a_i\}$ (which are the singular values of KX) are decreasing. Notice that this has the same form as the LS solution, except that the $\{x^{(i)}\}$ have replaced the $\{v^{(i)}\}$.

(d) KX : this is the truncated expansion method, using any set of vectors $\{x^{(i)}\}_1^k$ - one asks for the best solution of $Kf = g$ using vectors of the form $f = \sum_1^k c_i x^{(i)}$. Thus we find c as $\min_c \|KXc - g\|_2^2$, solving using the QR method or normal equations on KX . In particular, using the $x^{(i)}$ as above in (c) would give $c_i = \beta_i / a_i$, in which case this method would bear the same relation to MLS as SVD does to LS.

This method can be particularly useful if we can restrict the form of the solution f to such an expansion - again this is adding more constraints to the problem than in the other methods, but may result in there being a "unique" solution.

4. NUMERICAL EXAMPLES

All our examples stem from discretizations of the inverse Laplace transform:

$$\int_0^{\infty} e^{-st} f(t) dt = g(s).$$

This problem illustrates nicely all of the pitfalls associated with ill-posed problems, when different data $g(s)$ are used. The integral is discretized using n -point Gauss-Laguerre quadrature, so approximations are generated for $f(t)$ at the Gauss-Laguerre abscissae $\{t_j\}_1^n$. The sample points $\{s_i\}$ can vary, but normally n points equally spaced in $(0,n]$ gave the best results. Although all our examples come from known f and g , so that we can measure the "error" in f , this is rather artificial: in practical cases the data g will only be known at specific points $\{s_i\}$, and we can only measure the error in the residual sense ($\|Kf - g\|$).

For this problem, the singular vectors $v^{(i)}$ are asymptotic to zero (i.e. $v_j^{(i)} \rightarrow 0$ as $j \rightarrow n$), whereas the x -vectors $x^{(i)}$ are asymptotic to 1.0. As well, both sets of vectors satisfy the oscillation property: $v^{(i)}$ and $x^{(i)}$ each change sign $(i-1)$ times.

Example 1: $g(s) = \frac{1}{s+0.5}$, $f(t) = e^{-t/2}$.

Here we used $n=10$ points with $\{s_i\}_1^n$ equally spaced in $[1,10]$. For SVD, it is instructive to see the $\{\sigma_i\}$ and $\{\beta_i\}$ explicitly:

σ_i	.74	.28	.055	$.29 \times 10^{-2}$	$.17 \times 10^{-4}$	$.57 \times 10^{-8}$
β_i	-.92	-.018	-.014	$.61 \times 10^{-2}$	$-.16 \times 10^{-2}$	$.57 \times 10^{-2}$

σ_i	$.51 \times 10^{-3}$	$.52 \times 10^{-20}$	$.36 \times 10^{-25}$	$.99 \times 10^{-29}$
β_i	$.47 \times 10^{-2}$	$-.34 \times 10^{-2}$	$.19 \times 10^{-2}$	$-.12 \times 10^{-2}$

Notice that although no noise level ϵ is known, the $|\beta_i|$ decrease to about 10^{-2} while the σ_i decrease to zero. Clearly we should take $k = 3$; this gives a residual of .012 and a maximum error of 0.1. For the LS and MLS methods, it is not so clear how to choose α . Indeed it appears that for a large range of α , we can obtain reasonable solutions, even though they may be very different. This is particularly true of MLS, where the expansion vectors $x^{(i)}$ are asymptotic to 1, not 0, and the asymptotic nature of the solution to the inverse Laplace transform is not well-determined by the data $g(s_i)$. We give the basic results in Table 1 and the graphical results in Figure 1 (for SVD and LS) and Figure 2 (for MLS).

Table 1

α	LS		MLS	
	$ Kf-g $	$\max f_i - f(t_i) $	$ Kf-g $	$\max f_i - f(t_i) $
.1	.023	.083	.017	.40
.01	.012	.17	.012	.22
.001	.010	1.7	.012	1.6

Example 2: $g(s) = \frac{1}{s} - \frac{1}{s+0.5}$, $f(t) = 1 - e^{-t/2}$

Again we used 10 points with $\{s_i\}$ equally spaced in $[1,10]$. Of course the $\{\sigma_i\}$ are as in Example 1, and the $\{\beta_i\}$ are as follows:

β_i	-.31	-.16	-.034	$-.32 \times 10^{-2}$	$.81 \times 10^{-3}$	$-.57 \times 10^{-3}$
	$-.34 \times 10^{-3}$	$.17 \times 10^{-3}$	$-.59 \times 10^{-4}$	$.49 \times 10^{-4}$		

In this case the $\{\beta_i\}$ decay more rapidly to about 10^{-4} , so that it is appropriate to take $k=4$ terms (or possibly $k=3$); this gives a residual of .0011, and an error of 1.0 because the solution $f(t)$ is asymptotic to 1.0 (not 0 as in Example 1). However the SVD solution is still reasonable from the point of view of our definition, and it can only be seen as incorrect if more information is supplied about the problem.

For LS and MLS, the same comments apply as in Example 1. Here, although the MLS solution may look better (for some α) because of its asymptotic nature, there is no way to guarantee this for a given practical problem, again unless more is specified about the problem. We take the view that all solutions given here are reasonable. We again give LS and MLS results in Table 2 and the graphical solutions in Figures 3 and 4.

Table 2

α	LS		MLS	
	$\ Kf-g\ $	$\max f_i - f(t_i) $	$\ Kf-g\ $	$\max f_i - f(t_i) $
.1	.032	1.0	.013	.39
.01	.0033	1.0	.0017	.24
.001	.0011	1.0	.0011	.22

Example 3: $g(s) = \frac{2}{(s+0.5)^3}, \quad f(t) = t^2 e^{-t/2}.$

Here we used 20 points with the $\{s_i\}$ equally spaced in $[1,10]$. Truncating the SVD expansion at $k=7$ terms gave a residual of $.25 \times 10^{-4}$ and an error of 0.17. The LS and MLS solutions (for various α) are given in Table 3, with plots in Figures 5 and 6.

Table 3

α	LS		MLS	
	$\ Kf-g\ $	$\max f_i - f(t_i) $	$\ Kf-g\ $	$\max f_i - f(t_i) $
10^{-2}	$.60 \times 10^{-2}$	.54	$.27 \times 10^{-2}$	1.8
10^{-3}	$.28 \times 10^{-3}$	.30	$.27 \times 10^{-3}$	1.3
10^{-4}	$.35 \times 10^{-4}$	.17	$.32 \times 10^{-4}$	1.0
10^{-5}	$.21 \times 10^{-4}$	.71	$.24 \times 10^{-4}$	.29

Example 4: $g(s) = \frac{e^{-2s}}{s}, \quad f(t) = \begin{cases} 0, & t \leq 2 \\ 1, & t > 2. \end{cases}$

Here is one example of a discontinuous transform. The best SVD solution (with $k=4$) had a residual of $.79 \times 10^{-4}$ and an error of 1.0. The LS and MLS solutions are given below in Table 4, and graphically in Figures 7 and 8.

Table 4

α	LS		MLS	
	$ Kf-g $	$\max f_i - f(t_i) $	$ Kf-g $	$\max f_i - f(t_i) $
.1	.03	1.0	.021	.65
.01	$.38 \times 10^{-2}$	1.0	$.23 \times 10^{-2}$	.45
.001	$.42 \times 10^{-3}$	1.0	$.33 \times 10^{-3}$	.34
.0001	$.77 \times 10^{-4}$	1.0	$.99 \times 10^{-4}$	.31

Of course, the SVD and LS solutions are asymptotic to zero, and again the MLS solutions can be asymptotic to almost anything, depending upon the value of α . We feel that the most important point here is that all the solutions given here are reasonable, and can only be specified more precisely if more constraints are put on the problem.

5. CHOICE OF THE FREE PARAMETER

As we indicated earlier, the choice of k in the SVD method is fairly clear from the decay of the $\{\beta_i\}$ and $\{\sigma_i\}$. Unfortunately this does not appear to be the case for choosing α in the LS and MLS methods. Various strategies have been put forward for specifying α , notably the technique of generalized cross-validation (GCV), which has been successfully used for problems with non-smooth kernels (see Wahba et al [1979]). Here α is chosen to minimize the function

$$V(\alpha) = \frac{\frac{1}{n} \|g - Kf_{\alpha}\|_2^2}{\left(\frac{1}{n} \text{tr}(I - A(\alpha))\right)^2}$$

where $A(\alpha) = K(K^T K + \alpha^2 L^T L)^{-1} K^T$. Except for the $1/n$ factors, the numerator is just the square of the residual, and when $L=I$ (i.e. LS method) the denominator can be expressed as $\left(\sum_{i=1}^n \frac{\alpha^2}{\alpha^2 + \sigma_i^2}\right)$. In this case, notice that when some σ_i are very small (as for our smooth kernel) the corresponding terms in this sum will be very close to 1.0 for large ranges of α ; as well we have found that the residual is also nearly constant for large ranges of α , so that $V(\alpha)$ will be very flat and it will be very difficult to minimize. As an example of this, we plotted $V(\alpha)$ on a log-log scale for $0 \geq \log \alpha \geq -20$ for Example 1 above, and present the results in Figure 9. Similar results were obtained for the other examples.

Thus it seems impossible to choose α using this technique. However we feel this is not due to any intrinsic fault with the technique, but because reasonable solutions in our sense can be obtained over a wide range of the of the parameter α , even though they may look very different.

REFERENCES

- Bjorck and Elden (1979): Methods in numerical algebra for ill-posed problems. International Symposium on Ill-Posed Problems: Theory and Practice, University of Delaware.
- Golub, Heath, and Wahba (1979): Generalized Cross-Validation as a method for choosing a good ridge parameter. Technometrics ...
- C. van Loan (1976): Generalizing the singular value decomposition. SIAM J. Num. Ana. 13, 76-83.
- J. Varah (1979): A practical examination of some numerical methods for linear discrete ill-posed problems. SIAM Review 21, 100-111.

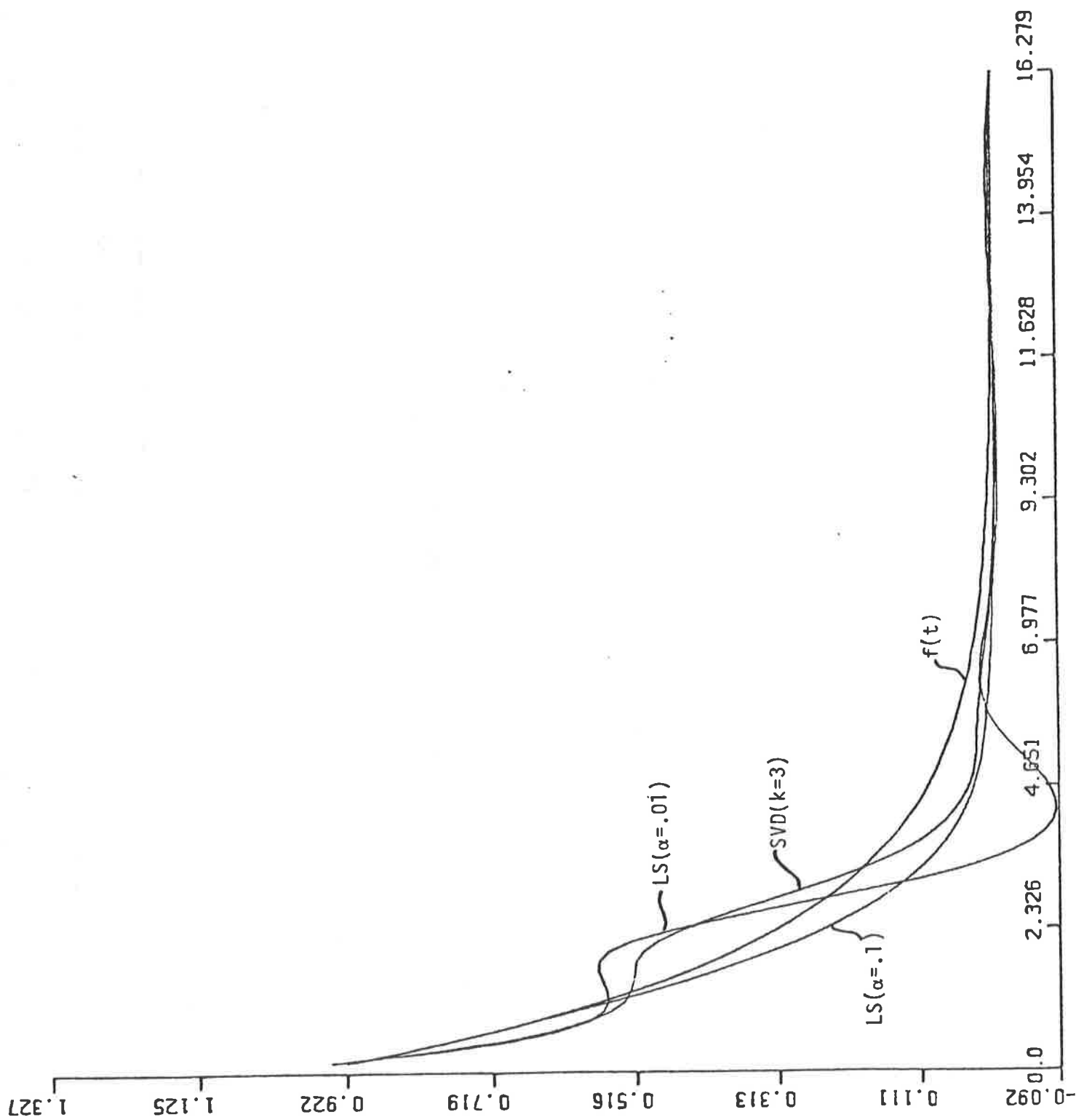


Figure 1: SVD and LS solutions for Example 1.

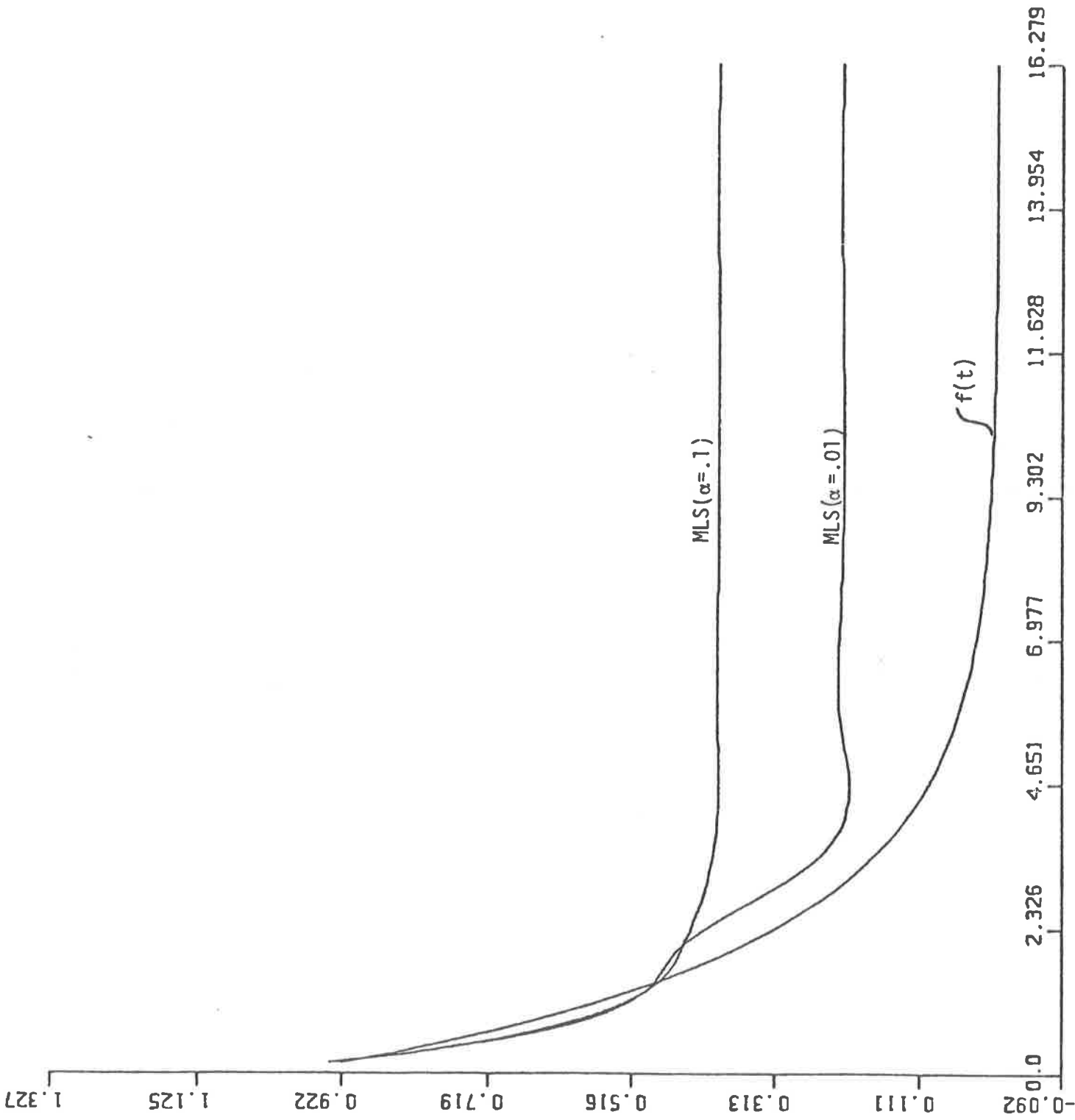


Figure 2: MLS solutions for Example 1.

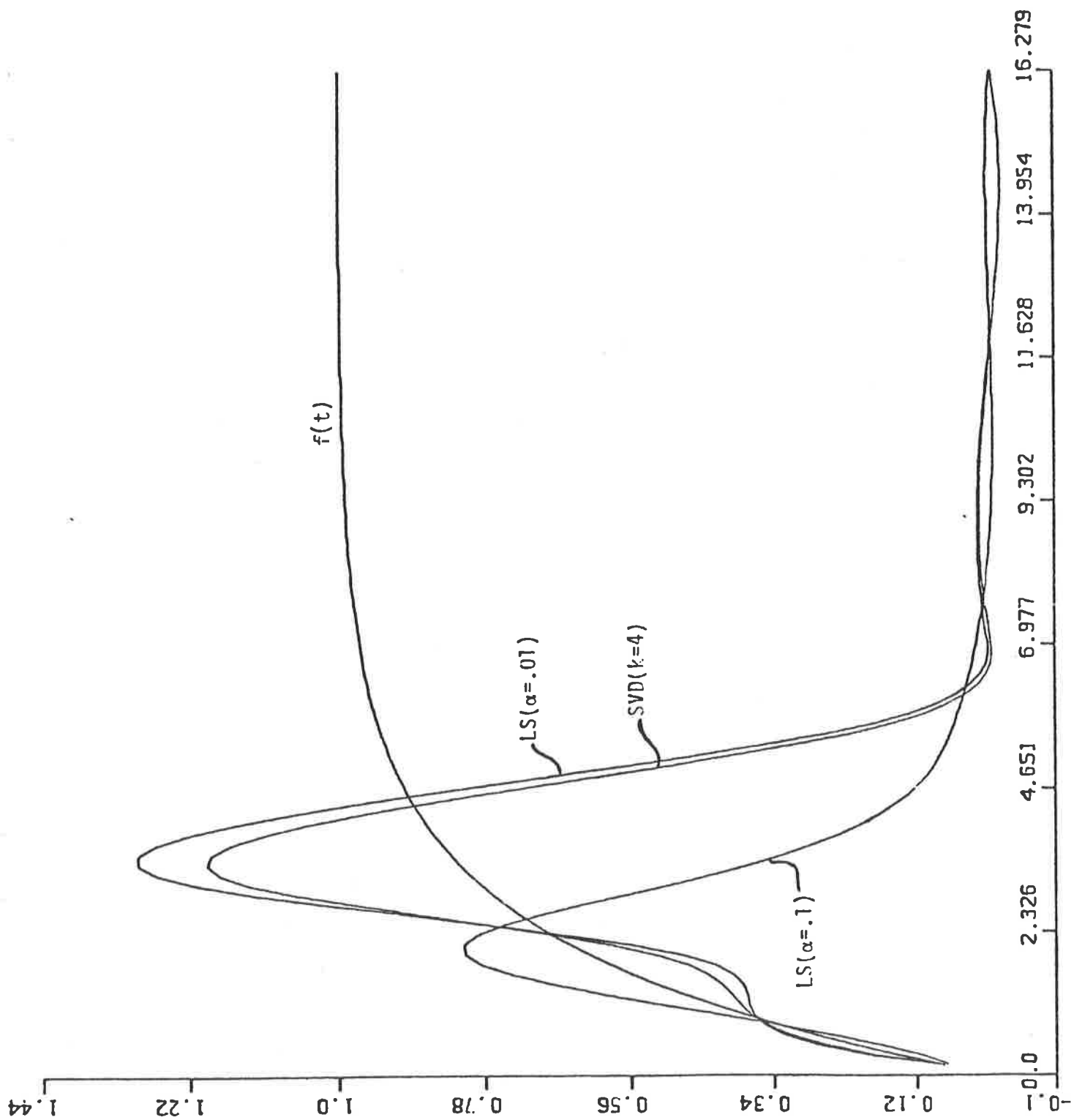


Figure 3: SVD and LS solutions for Example 3.

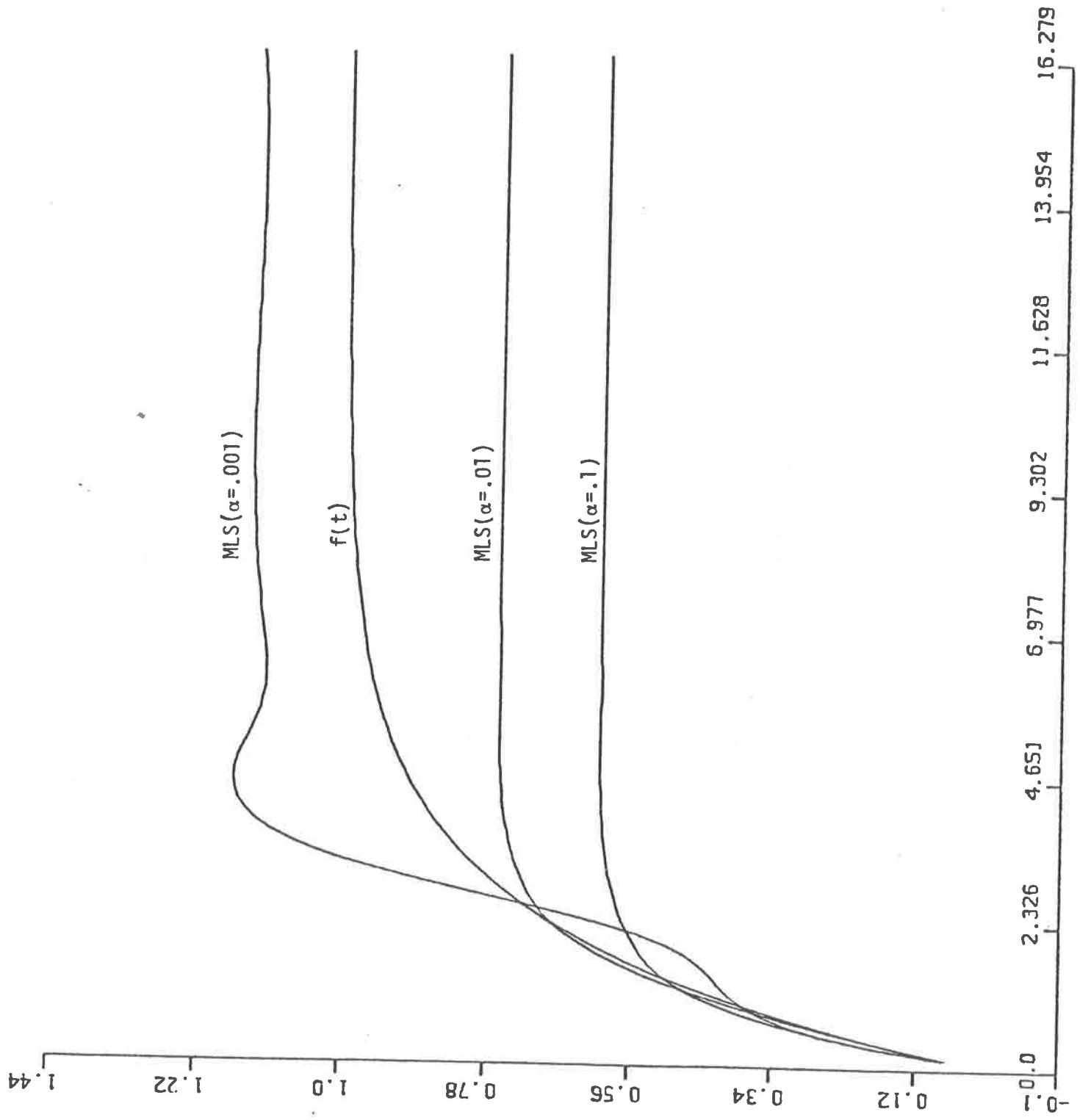


Figure 4: MLS solutions for Example 2.

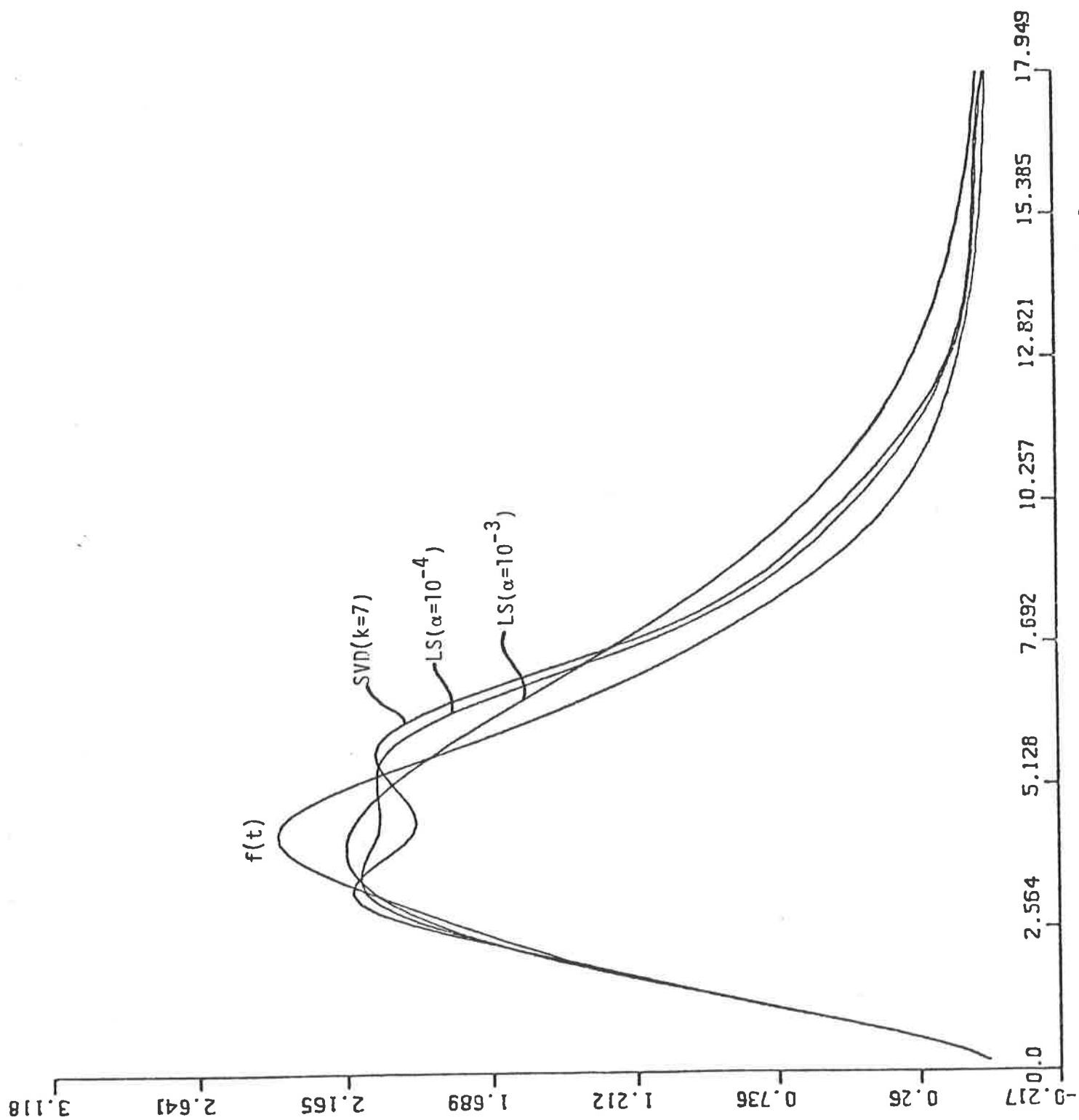


Figure 5: SVD and LS solutions for Example 3.

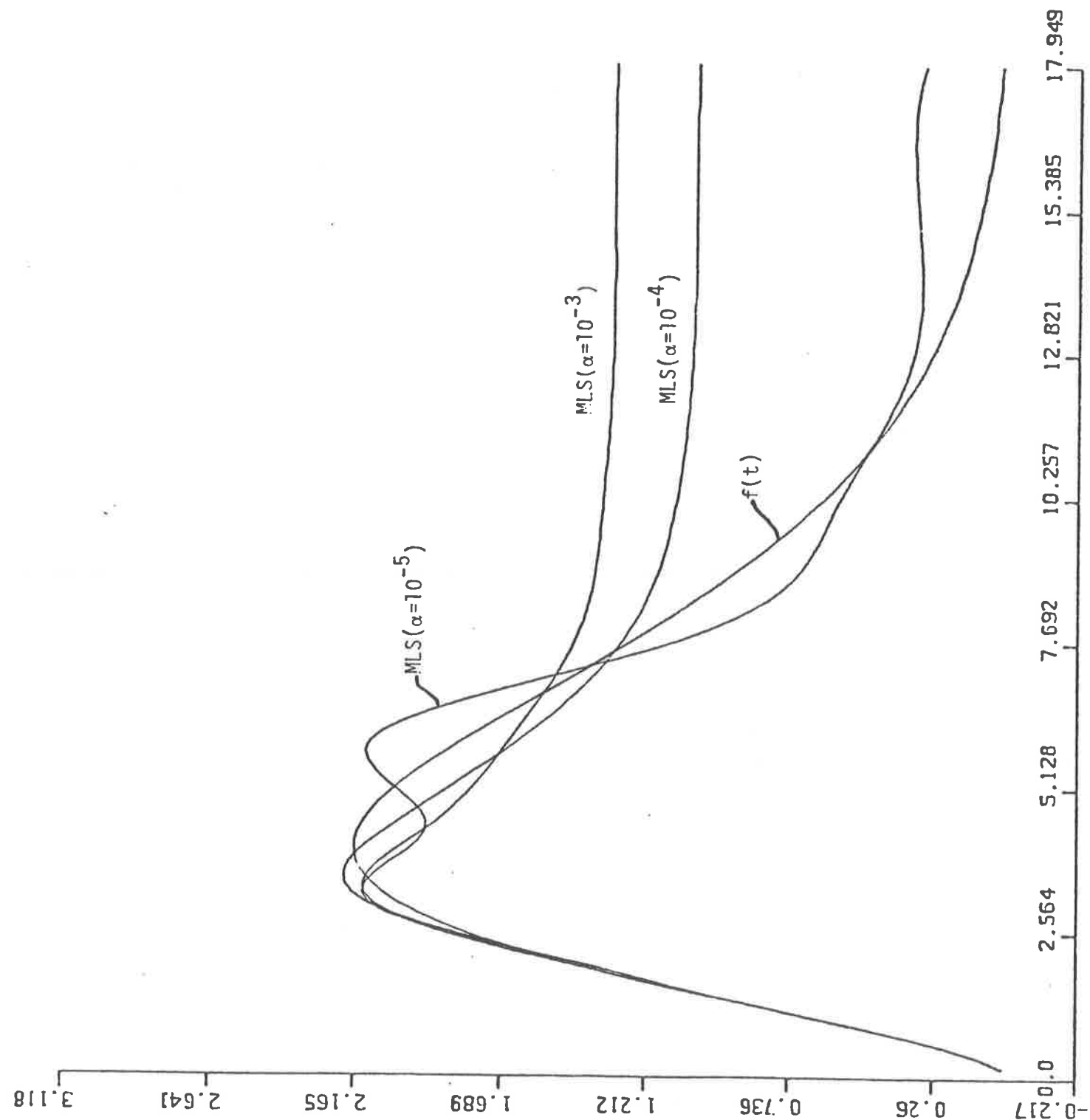


Figure 6: MLS solutions for Example 3.

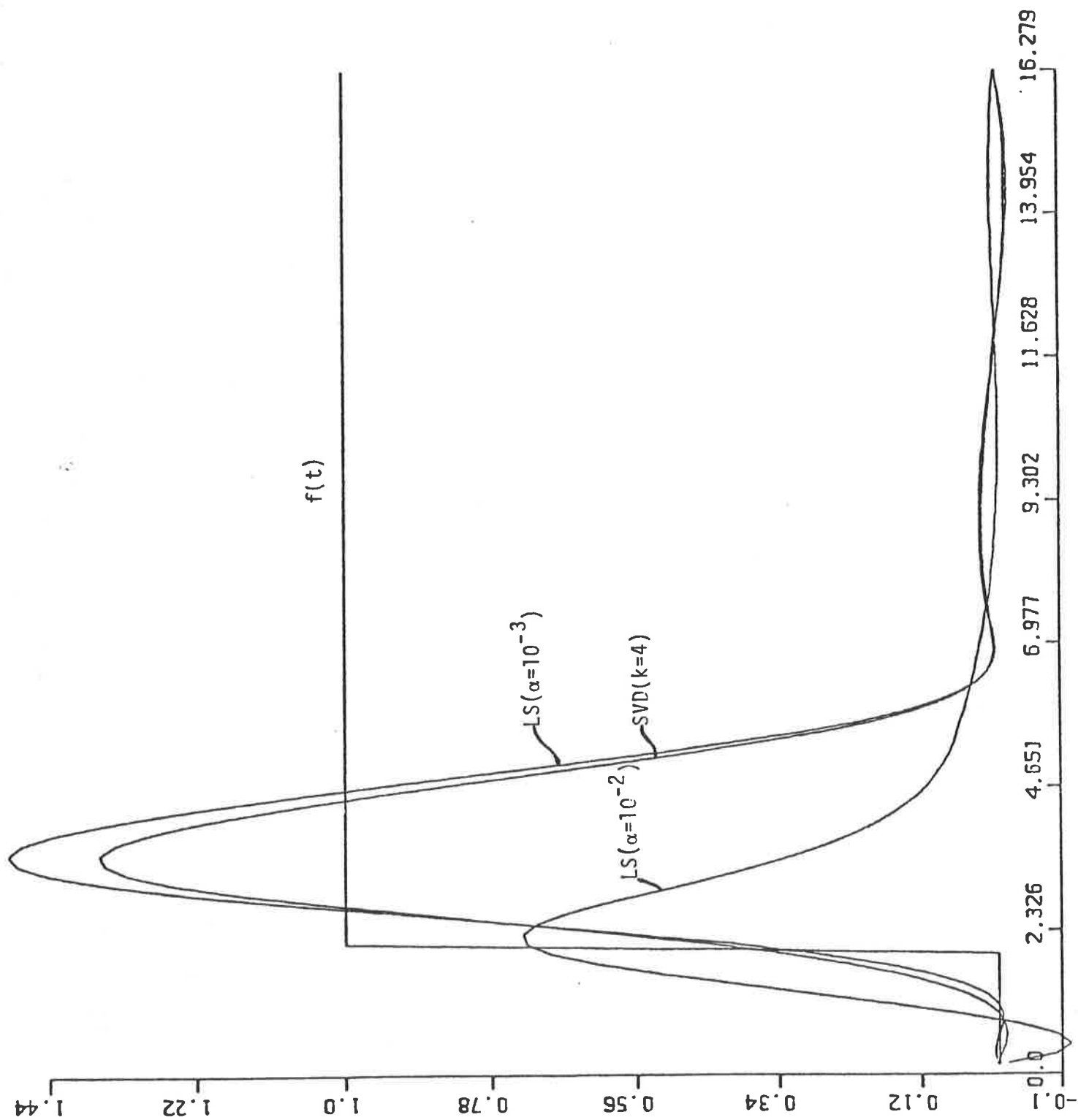


Figure 7: SVD and LS solutions for Example 4.

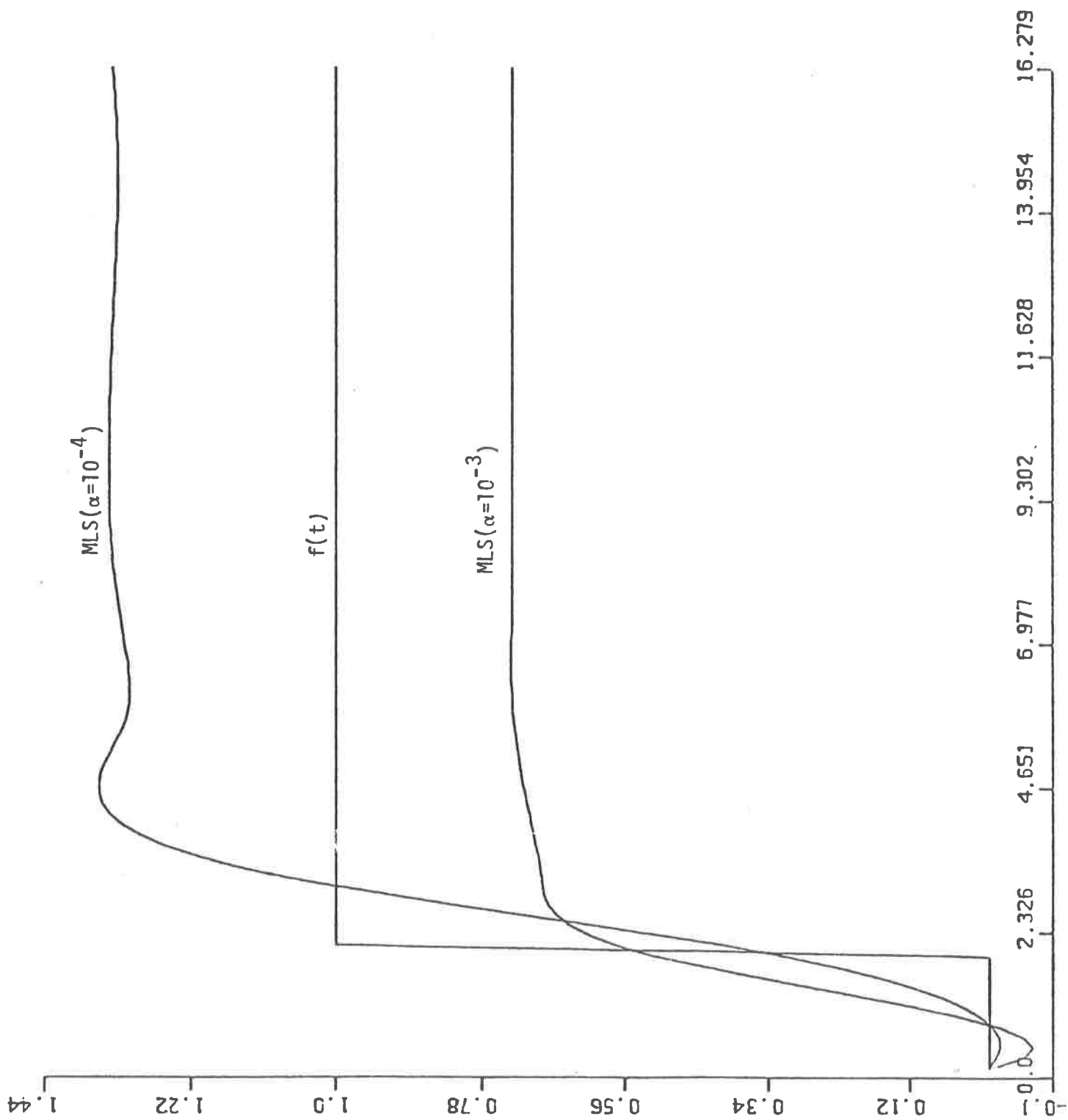


Figure 8: MLS solutions for Example 4.

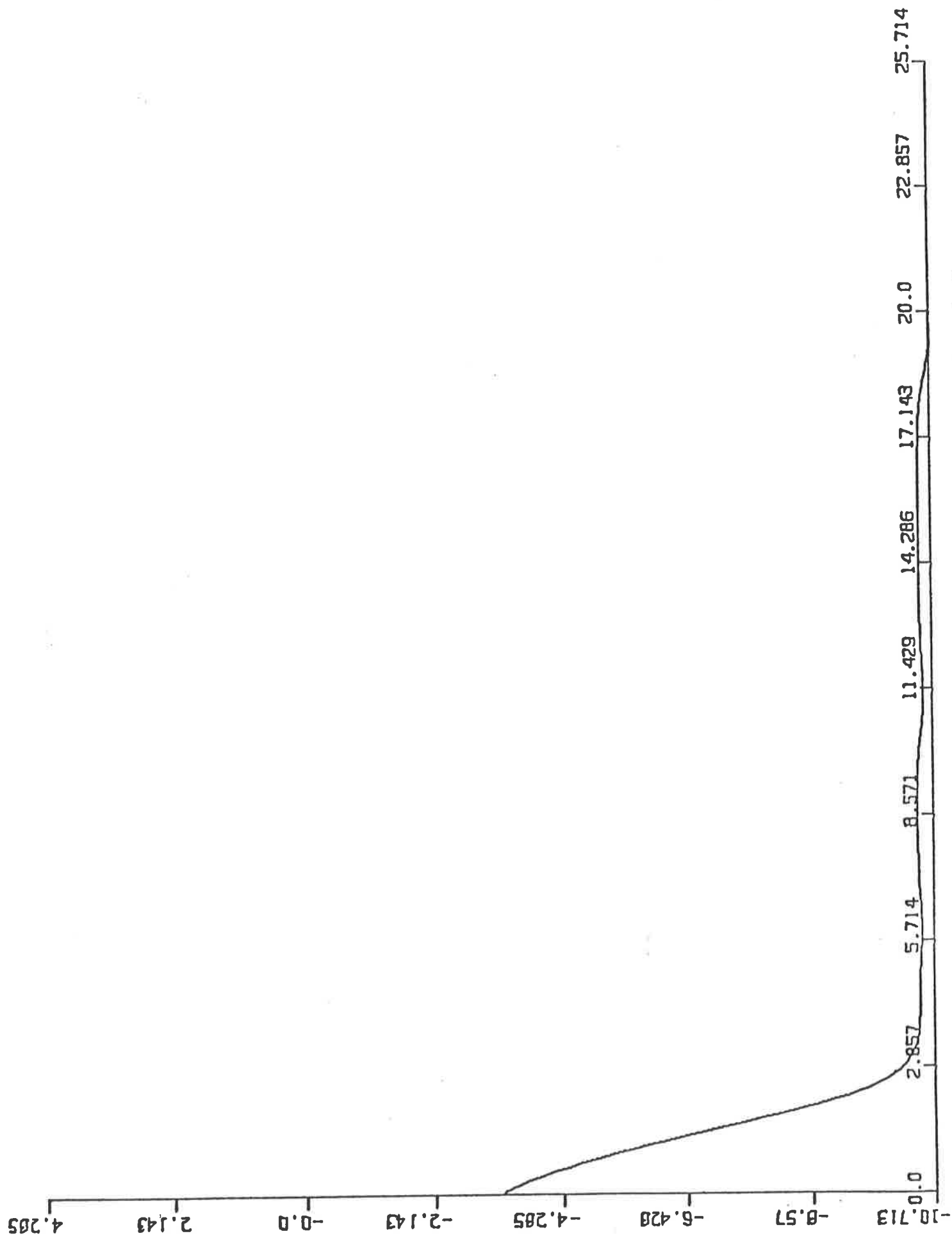


Figure 9. log-log plot of GCV function for Example 1.

