Mathematical and Computational Sciences Division

NISTIR 6131

Information Technology Laboratory

TRUNCATING THE SINGULAR VALUE DECOMPOSITION FOR ILL-POSED PROBLEMS

Bert W. Rust

July 1998

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards and Technology
Gaithersburg, MD 20899

TRUNCATING THE SINGULAR VALUE DECOMPOSITION FOR ILL-POSED PROBLEMS

Bert W. Rust
Mathematical and Computational Sciences Division
Building 820, Room 365
National Institute of Standards and Technology
Gaithersburg, MD 20899

Abstract

Discretizing the first-kind integral equations which model many physical measurement processes yields an ill-conditioned linear regression model $\mathbf{b} = \mathbf{A}\mathbf{x}^* + \boldsymbol{\eta}$, where $\mathbf{x}^*$ is a vector representation of the function being measured, $\mathbf{A}$ is an instrument response matrix, $\mathbf{b}$ is a vector of measurements, and $\boldsymbol{\eta}$ is a vector of unknown, random measuring errors. Least squares estimation usually gives a sum of squared residuals much smaller than the expected value and a wildly oscillating, physically implausible estimate of $\mathbf{x}^*$. These symptoms suggest that the least squares estimate captures part of the variance that properly belongs in the residuals. One strategy for shifting some of this variance to the residuals and simultaneously stabilizing the estimate is to truncate the singular value decomposition $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\boldsymbol{\Sigma}$ is a diagonal matrix of singular values. All of the singular values below some threshold value are reset to zero to give a new matrix $\boldsymbol{\Sigma}_{tr}$, and the estimated solution is calculated from the generalized inverse of the matrix $\mathbf{U}\boldsymbol{\Sigma}_{tr}\mathbf{V}^T$. The most delicate part of this procedure is the determination of the truncation threshold. Conventionally this has been regarded as a problem of determining the "numerical rank" of $\mathbf{A}$, but in most cases $\mathbf{A}$ is clearly not rank-deficient. This paper suggests an alternate strategy which uses the variances of the measuring errors to specify a truncation for the elements of the rotated measurement vector $\mathbf{U}^T\mathbf{b}$. The idea is to zero all of the components that are dominated by the measurement errors and compute the estimate using the full rank matrix. The problem of setting the truncation threshold becomes one of deciding whether or not a measured value is significantly different from zero, a procedure familiar to most experimentalists. The paper also develops some new diagnostics for the residuals which are useful not only for choosing the truncation level for the $(\mathbf{U}^T\mathbf{b})_i$, but also for assessing the quality of an estimate obtained by any procedure.

1 First Kind Integral Equations with Uncertainties

First kind integral equations,

$$y(t) = \int_a^b K(t,\xi)x(\xi)\, d\xi \ , \tag{1.1}$$

where $K(t,\xi)$ and $y(t)$ are known functions, are routinely used to model instrument distortions in measuring an unknown function $x(\xi)$. In that context, they are usually written as a system of equations

$$y_i = \int_a^b K_i(\xi)x(\xi)\, d\xi + \epsilon_i \ , \quad i = 1, 2, \ldots, m \ , \tag{1.2}$$

where the $K_i(\xi)$ are known (previously measured or calculated) response functions of the instrument, the $y_i \cong y(t_i)$ are measured values, corresponding to a discrete mesh $t_1, t_2, \ldots, t_m$, and the ϵ_i are random, zero-mean measuring errors. In order to estimate $x(\xi)$ it is necessary to further discretize the system, in the process replacing it with a linear regression model

$$\mathbf{y} = \mathbf{K}\mathbf{x}^* + \boldsymbol{\epsilon} \ , \tag{1.3}$$

where $\mathbf{y}$ is the m-vector of measurements, $\mathbf{K}$ is a known $m \times n$ matrix, with $m \geq n$, and $\mathbf{x}^*$ is an unknown n-vector whose components are either discrete point estimates of $x(\xi)$ on some mesh $\xi_1, \xi_2, \ldots, \xi_n$, or are the unknown coefficients in a truncated expansion of $x(\xi)$ in terms of some set of basis functions spanning the space of possible solutions. The vector $\boldsymbol{\epsilon}$ is an m-vector of random measuring errors satisfying

$$\mathcal{E}(\boldsymbol{\epsilon}) = \mathbf{0} \ , \qquad \mathcal{E}\left(\boldsymbol{\epsilon}\, \boldsymbol{\epsilon}^T\right) = \mathbf{S}^2 \ , \tag{1.4}$$

where $\mathcal{E}$ is the expectation operator, $\mathbf{0}$ is the m-dimensional zero vector and $\mathbf{S}^2$ is the positive definite variance-covariance matrix for $\boldsymbol{\epsilon}$. In most problems the measurement errors are statistically independent so

$$\mathbf{S}^2 = \mathbf{diag}(s_1^2,\ s_2^2,\ \ldots,\ s_m^2) \ , \tag{1.5}$$

where $s_1, s_2, \ldots, s_m$ are the standard deviations of the errors. If $\mathbf{S}^2$ is not diagonal, the model can be transformed into one with a diagonal variance matrix by premultiplying (1.3) by the inverse of the lower triangular Cholesky factor of $\mathbf{S}^2$.

There seems to be a general misapprehension in the numerical analysis community that estimates of the s_i are seldom available, but in fact good experimenters routinely provide them. Estimates of the measurement errors are not considered to be something extra, but rather are an integral part of the measurements, and published graphs of measured data will usually report them as $\pm 1\sigma$ error bars on the plotted points. An analyst who fails to use this information implicitly assumes that $\mathbf{S}^2 = s^2\mathbf{I}_m$ where $\mathbf{I}_m$ is the m-th order identity matrix and s is an unknown scalar that can be, but usually is not, estimated from the sum of squared residuals for the least squares solution. Such an assumption may be tenable if all the y_i have roughly the same magnitude, but in most cases, they span a range of values, and the magnitudes of the s_i vary (usually nonlinearly) with the magnitudes of the y_i.

In the following it will be assumed that $\mathbf{S}$ is a known diagonal matrix. It will also be assumed, as is often the case, that the errors are samples from a multivariate normal distribution, i.e., that $\boldsymbol{\epsilon} \sim N(\ \mathbf{0}\ ,\ \mathbf{S}^2\ )$. In §2, these assumptions will be used to rescale the problem to have errors with a standard normal distribution and to derive a statistical diagnostic for the sum of squared residuals for any estimate of the solution vector. In §3 a variant of the well known Phillips problem [16] is subjected to this scaling and the ordinary least squares estimate is calculated and found to be unsatisfactory. In §4 the residuals for that estimate are subjected to diagnostic tests based on the expected value of their sum of squares and on their distribution when considered as a time-series. The diagnostics that are developed there are useful for testing any estimate, no matter how it may be obtained, and will be used throughout the remainder of the paper.

The singular value decomposition is introduced in §5 and the conventional method for stabilizing solution estimates by truncating the distribution of singular values is described in §6. This method is based on an assumption that the matrix is rank deficient and depends critically on the proper determination of its "numerical rank." But for most ill-posed problems, the matrix is demonstrably not rank deficient so there is no good reason for setting any of the singular values to zero. An alternate approach to truncating the decomposition is given in §7. Rather than zeroing some of the singular values, one instead zeroes small elements of the rotated right hand side vector formed by premultiplying the vector of measurements by the transpose of the matrix of left singular vectors, i.e., the leftmost factor in the singular value decomposition. The errors in this rotated measurement vector have a standard normal distribution, so it is possible to establish a statistical criterion for determining whether or not a given element is significantly different from zero. The new method is applied to the modified Phillips problem to obtain satisfactory estimates for 3 different truncation levels and the optimum level is taken to be the one which best satisfies the diagnostic tests described in §4. In §8 the method is successfully applied to real-world measurements of the energy spectrum of neutrons produced by a certain nuclear reaction. Finally, §9 gives a brief discussion of algorithmic considerations and of how the method can be extended when the knowledge of the measurement errors is not as complete as might be desired.

2 Scaling the Problem

The linear regression model in the preceding section can be written

$$\mathbf{y} = \mathbf{K}\mathbf{x}^* + \boldsymbol{\epsilon} \, , \quad \boldsymbol{\epsilon} \sim N(\, \mathbf{0} \, , \, \mathbf{S}^2 \,) \, , \tag{2.1}$$

but it is advantageous to scale it with the matrix $\mathbf{S}^{-1}$. Let

$$\mathbf{b} \equiv \mathbf{S}^{-1}\mathbf{y} \, , \quad \mathbf{A} \equiv \mathbf{S}^{-1}\mathbf{K} \, , \quad \boldsymbol{\eta} \equiv \mathbf{S}^{-1}\boldsymbol{\epsilon} \, , \tag{2.2}$$

and note that by a standard theorem of multivariate statistics [2, Thm. 2.4.4], $\boldsymbol{\epsilon} \sim N(\, \mathbf{0} \, , \, \mathbf{S}^2 \,)$ implies that $\boldsymbol{\eta} \sim N(\, \mathbf{S}^{-1}\mathbf{0} \, , \, \mathbf{S}^{-1}\mathbf{S}^2[\mathbf{S}^{-1}]^T \,)$, so the scaled model can be written

$$\mathbf{b} = \mathbf{A}\mathbf{x}^* + \boldsymbol{\eta} \, , \quad \boldsymbol{\eta} \sim N(\, \mathbf{0} \, , \, \mathbf{I}_m \,) \, , \tag{2.3}$$

or

$$\mathbf{b} \sim N(\, \mathbf{A}\mathbf{x}^* \, , \, \mathbf{I}_m \,) \, . \tag{2.4}$$

To see the advantage of this scaling, let $\hat{\mathbf{x}}$ be an estimate of $\mathbf{x}^*$ and

$$\hat{\mathbf{r}} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} \, , \tag{2.5}$$

be the corresponding residual vector. Since the regression model can also be written

$$\boldsymbol{\eta} = \mathbf{b} - \mathbf{A}\mathbf{x}^* \, , \tag{2.6}$$

it is clear that *an estimate* $\hat{\mathbf{x}}$ *is acceptable only if* $\hat{\mathbf{r}}$ *is a plausible sample from the* $\boldsymbol{\eta}$*-distribution.*

Since $\mathbf{b} - \mathbf{A}\mathbf{x}^* \sim N(\, \mathbf{0} \, , \, \mathbf{I}_m \,)$, it follows from another standard statistical theorem [14, page 140] that

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2 \equiv (\mathbf{b} - \mathbf{A}\mathbf{x}^*)^T(\mathbf{b} - \mathbf{A}\mathbf{x}^*) \sim \chi^2(m) \, , \tag{2.7}$$

where $\chi^2(m)$ denotes the Chi-squared distribution with m degrees of freedom, and hence that

$$\mathcal{E}\left\{\|\mathbf{b}-\mathbf{A}\mathbf{x}^*\|^2\right\} = m\ , \qquad \mathrm{Var}\left\{\|\mathbf{b}-\mathbf{A}\mathbf{x}^*\|^2\right\} = 2m\ . \tag{2.8}$$

These two quantities provide rough bounds for the sum of squared residuals that might be expected from a reasonable estimate of $\mathbf{x}^*$. An estimate that gives

$$m - \sqrt{2m} \le \|\mathbf{b}-\mathbf{A}\hat{\mathbf{x}}\|^2 \le m + \sqrt{2m} \tag{2.9}$$

would be quite reasonable, but any $\hat{\mathbf{x}}$ whose sum of squared residuals falls outside the interval $\left[m - 2\sqrt{2m}, m + 2\sqrt{2m}\right]$ would be suspect. These rough indicators can be sharpened considerably by using the cumulative distribution function for $\chi^2(m)$. More details on this point are given in Appendix A.

3 Standard Linear Regression

The standard approach for the linear regression problem defined by [2.3] is to assume that $\mathrm{rank}(\mathbf{A}) = n$ and seek the minimum variance, unbiased estimator for $\mathbf{x}^*$ by solving the least squares minimization problem

$$r^2_{\min} = \min_{\mathbf{x}\in R^n} \left\{(\mathbf{b}-\mathbf{A}\mathbf{x})^T(\mathbf{b}-\mathbf{A}\mathbf{x})\right\}\ . \tag{3.1}$$

The solution,

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}\ , \tag{3.2}$$

is called the *best linear unbiased estimate* of $\mathbf{x}^*$, but it is well known [18, Chapt. 1], [19, Chapt. 6], [21, Chapt. 2] that for regression models obtained by discretizing first kind integral equations, the elements of $\hat{\mathbf{x}}$ are pathologically sensitive to small variations in the elements of $\mathbf{b}$, so the presence of the measuring errors leads to estimates that are totally unphysical, typically oscillating wildly between extreme positive and negative values.

A useful test problem, with known solution, which shares many of the characteristics of real instrument correction problems can be obtained by discretizing a variant of the well known Phillips equation [16] and adding some random errors to the discrete y_i. The problem and a discretization with $m = 300$ and $n = 241$ are described in detail in Appendices B and C. The functions $y(t)$ and $x(\xi)$ are plotted on the right in Figure 1, and on the left are plotted the functions $K(t,\xi_j)$ for the discrete values $\xi_j = -3.0\,,\ -1.5\,,\ 0\,,\ 1.5\,,\ 3.0\,.$ All of the 241 $K(t,\xi_k)$ have the same shape and subtend unit area. The 300 standard deviations s_i range in value from 3.49×10^{-11} to 7.78×10^{-6} so the errors in

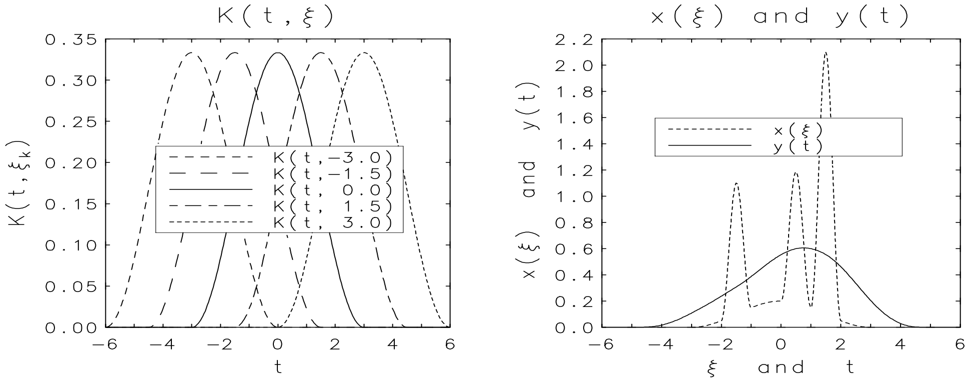

Figure 1: A variant of the Phillips problem

the y_i are much smaller than the thickness of the curve in the plot of $y(t)$. But these small errors produce large oscillations in the least squares estimate of the solution vector. This is shown in the plot in the upper left corner of Figure 2 where the barely discernable dashed curve is the true solution $\mathbf{x}^*$, and the solid curve oscillating wildly about it is the least squares estimate $\hat{\mathbf{x}}$. The magnitudes of these oscillation are roughly 10^7 times greater than the largest random errors in the y_i. Such a large amplification of the measuring errors is the rule rather than the exception for ill posed problems.

4 Estimate Diagnostics

The most commonly used estimate diagnostic for linear regression problems is the sum of squared residuals. For the problem in the preceding section the least squares estimate gives

$$r_{\min}^2 = \min_{\mathbf{x} \in R^n} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|^2 = 43.01 \tag{4.1}$$

which is much smaller than the value expected for the true solution which, by Eq. (2.8), is

$$\mathcal{E}\left\{\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2\right\} = 300 \; . \tag{4.2}$$

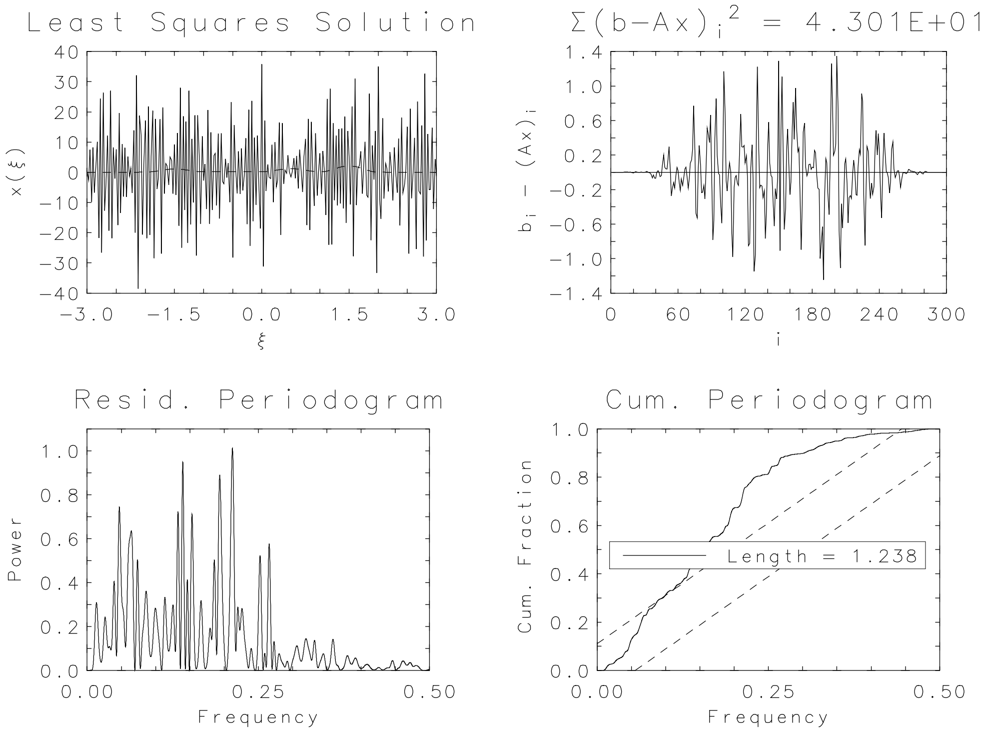

Figure 2: Least squares estimate and diagnostics for the modified Phillips problem

From Eq. (2.8) it also follows that the standard deviation for this expected value is $\sqrt{600} = 24.49$, so the least squares $r^2_{\min}$ is more than 10 standard deviations smaller than the expected value. Thus, without even inspecting the estimate itself, one could argue that $\hat{\mathbf{x}}$ is unlikely to be a good approximation to $\mathbf{x}^*$.

Another good diagnostic for judging the acceptability of an estimate can be gotten by considering the elements of the vectors $\boldsymbol{\eta}$ and $\hat{\mathbf{r}}$ as time series, with the element number taken as the time variable. Since the η_i are statistically independent and identically normally distributed, they form a normally distributed white noise series, so the residuals for an acceptable estimate should constitute a realization of such a series. The residuals for the least squares estimate are plotted against element number at the upper right of Figure 2. It is clear by inspection that they are not even approximately white noise, and this is not surprising since it can be shown [10, Section 6.9.1] that

$$\hat{\mathbf{r}} \sim N\left(\mathbf{0},\ \left[\mathbf{I}_m - \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\right]\right) \tag{4.3}$$

with the variance matrix becoming approximately diagonal only when m is much larger than n.

A rigorous test for white noise is based on the periodogram [6, Chapt. 7] which is an estimate of the power spectrum on the frequency interval $0 \leq f \leq \frac{1}{2T}$, where T is the sample spacing for the time variable. Here, the time variable is the element number i, so $T = 1$. The graph at the lower left of Figure 2 is a plot of a periodogram estimate at 4097 equally spaced frequency points on the interval $[0, 0.5]$. The details on how it was computed are given in Appendix D. For a white noise series, the variance would be distributed uniformly over the frequency range, so the fact that there is much more power on the interval $[0, 0.25]$ than on $[0.25, 0.5]$ suggests that the residuals should not be regarded as white noise. A formal test of this hypothesis uses the cumulative periodogram [6, Chapt. 7] (also described in Appendix D) which is plotted as a solid curve at the lower right of the figure. The legend box gives the length of this curve which can be compared with the value 1.11803 expected for a pure white noise spectrum. The two dashed lines enclose a 95% confidence band for white noise. The cumulative periodogram ordinates for such a series should lie outside this band for at most 5% of the frequencies. Since 2803 of the 4096 calculated ordinates lie outside the band, it is safe to conclude that the least squares residuals are not white noise.

5 The Singular Value Decomposition

Insight into the failure of the least squares method to give a reasonable estimate of $\mathbf{x}^*$ can be gotten by substituting the singular value decompostion of $\mathbf{A}$ into the model (2.3). Let

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{O} \end{pmatrix} \mathbf{V}^T , \quad \mathbf{\Sigma} = \mathbf{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) , \tag{5.1}$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, and

$$\mathbf{U}^T\mathbf{U} = \mathbf{I}_m = \mathbf{U}\mathbf{U}^T , \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}_n = \mathbf{V}\mathbf{V}^T . \tag{5.2}$$

Substituting (5.1) into (2.3) and premultiplying by $\mathbf{U}^T$ gives

$$\mathbf{U}^T\mathbf{b} = \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{O} \end{pmatrix} \mathbf{V}^T\mathbf{x}^* + \mathbf{U}^T\boldsymbol{\eta} , \quad \mathbf{U}^T\boldsymbol{\eta} \sim N(\,\mathbf{0}\,,\,\mathbf{I}_m\,) , \tag{5.3}$$

with the distribution of the $\mathbf{U}^T\boldsymbol{\eta}$ vectors unchanged because premultiplication by an orthogonal matrix simply rotates all the vectors in the distribution through the same angle. Also, premultiplying a vector by an orthogonal matrix leaves its length unchanged so Eq. (2.7) can be replaced by

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2 = \left\| \mathbf{U}^T\mathbf{b} - \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{O} \end{pmatrix} \mathbf{V}^T\mathbf{x}^* \right\|^2 \sim \chi^2(m) \; , \tag{5.4}$$

with the $\chi^2(m)$ distribution following from the fact that the $(\mathbf{U}^T\boldsymbol{\eta})_i \sim n(0,1)$, idependently. Partitioning the $m \times m$ matrix $\mathbf{U}$,

$$\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2) \; , \tag{5.5}$$

with $\mathbf{U}_1$ an $m \times n$ and $\mathbf{U}_2$ an $m \times (m-n)$ submatrix allows the least squares problem to be written

$$\min_{\mathbf{x}\in R^n} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \min_{\mathbf{x}\in R^n} \left\| \begin{pmatrix} \mathbf{U}_1^T \\ \mathbf{U}_2^T \end{pmatrix} \mathbf{b} - \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{O} \end{pmatrix} \mathbf{V}^T\mathbf{x} \right\|^2 \; , \tag{5.6}$$

from which it is quite evident that the least squares solution must satisfy

$$\mathbf{V}^T\hat{\mathbf{x}} = \mathbf{\Sigma}^{-1}\mathbf{U}_1^T\mathbf{b} \; , \tag{5.7}$$

and that the minimum sum of squared residuals is

$$r^2_{\text{min}} = \|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|^2 = \|\mathbf{U}_2^T\mathbf{b}\|^2 \; . \tag{5.8}$$

These last two equations can also be written

$$\left(\mathbf{V}^T\hat{\mathbf{x}}\right)_i = \frac{\left(\mathbf{U}^T\mathbf{b}\right)_i}{\sigma_i} \; , \quad i = 1, 2, \cdots, n \; , \tag{5.9}$$

and

$$r^2_{\text{min}} = \|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|^2 = \sum_{i=n+1}^{m} \left(\mathbf{U}^T\mathbf{b}\right)_i^2 \; , \tag{5.10}$$

and it is easy to see that Eq. (4.3) can be written

$$\hat{\mathbf{r}} \sim N\left(\mathbf{0}, \; \mathbf{I}_m - \mathbf{U}_1\mathbf{U}_1^T\right) \; . \tag{5.11}$$

For the modified Phillips problem the value of r^2_{min} was too small by 10 standard deviations which suggests that some of the $(\mathbf{U}^T\mathbf{b})_i$ values in the sequence (5.9) more properly belong in the sum in (5.10). This reasoning leads to the idea of truncating the singular value decomposition.

6 Truncating the Singular Value Distribution

Historically, the idea of truncating the SVD has been treated as a problem of determining the "numerical rank" of the matrix $\mathbf{A}$. All of the computed singular values smaller than some threshold value are treated as zeroes which were corrupted by rounding errors into small non-zero quantities. Thus if σ_p is the smallest singular value greater than the truncation threshold, then one replaces the matrix $\mathbf{\Sigma}$ by a truncated matrix

$$\mathbf{\Sigma}_{tr} = \mathbf{diag}(\sigma_1, \ldots, \sigma_p, 0, \ldots, 0) \tag{6.1}$$

whose generalized inverse is given by

$$\mathbf{\Sigma}_{tr}^{\dagger} = \mathbf{diag}\left(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_p}, 0, \ldots, 0\right) \; . \tag{6.2}$$

The value p is said to be the "numerical rank" of $\mathbf{A}$. If $\tilde{\mathbf{x}}$ is the estimate of the solution to the truncated problem, then it is required to satisfy

$$\mathbf{V}^T\tilde{\mathbf{x}} = \mathbf{\Sigma}_{tr}^{\dagger}\mathbf{U}_1^T\mathbf{b} \tag{6.3}$$

rather than Eq. (5.7), which means that Eqs. (5.9) and (5.10) are replaced by

$$\left(\mathbf{V}^T\tilde{\mathbf{x}}\right)_i = \begin{cases} \dfrac{\left(\mathbf{U}^T\mathbf{b}\right)_i}{\sigma_i} & , \quad i = 1, 2, \cdots, p \; , \\ 0 & , \quad i = p+1, \cdots, n \; , \end{cases} \tag{6.4}$$

and

$$\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|^2 = \sum_{i=p+1}^{m} \left(\mathbf{U}^T\mathbf{b}\right)_i^2 \; , \tag{6.5}$$

and Eq. (5.11) is replaced by

$$\tilde{\mathbf{r}} \sim N\left[\left(\mathbf{I}_m - \mathbf{U}_p\mathbf{U}_p^T\right)\mathbf{A}\mathbf{x}^* \, , \; \mathbf{I}_m - \mathbf{U}_p\mathbf{U}_p^T\right] \; , \tag{6.6}$$

where $\mathbf{U}_p$ is the submatrix formed by the first p columns of $\mathbf{U}$. Since $p < n$, it is clear that the variance matrix for $\tilde{\mathbf{r}}$ is more nearly diagonal than that for $\hat{\mathbf{r}}$, and that if $p \ll m$, then the $\tilde{r}_i$ will be approximately independently distributed.

The truncated singular value approach was first suggested by Golub and Kahan [8] who noted its similarity to the theoretical treatment given by Smithies [17, Chapt. 8] for the singular functions and singular values of first kind integral equations. One of the first to use the method was Richard Hanson [13] who suggested that the truncation threshold should be chosen to be the smallest integer p such that

$$\sum_{i=p+1}^{m} \left(\mathbf{U}^T\mathbf{b}\right)_i^2 < m \; . \tag{6.7}$$

In view of Eqs. (2.7) and (2.8), this seems a very sensible choice, but his suggestion was apparently ignored by most other workers in the field who were preoccupied with the problem of rigorously defining the notion of the "numerical rank" of the matrix. An especially influential paper in this line of research was a technical report [9] by Golub, Klema and Stewart who used techniques from perturbation analysis to define a very complicated criterion for rank determination which was norm-dependent and involved a triplet of numbers (δ, ϵ, p) with δ and ϵ being numbers satisfying $\sigma_p \geq \delta > \epsilon \geq \sigma_{p+1}$. The idea was to find a clear gap in the distribution of singular values and zero all those on the low side of the gap. The trouble with this approach is that matrices arising from discretizing first kind integral equations almost never display such a gap. This fact was clearly enunciated by P.C. Hansen [11] who proposed "... a natural division of ill-conditioned matrices into two classes: those with well-determined numerical rank and those with ill-determined numerical rank." By the latter he meant matrices whose singular values decayed smoothly from σ_1 to σ_n with no obvious gap between the larger and smaller ones. More recently, in his new book [12, Chapt. 4], he has written "The main feature of these problems is that all the singular values of the coefficient matrix decay gradually to zero, with no gap anywhere in the spectrum. ... and therefore the concept of 'numerical rank' is not useful for these problems." Nevertheless he devoted large segments of the book to problems of rank determination and rank revealing decompositions, and gave a good review of the very sizable literature on these subjects.

The singular values for the modified Phillips problem are plotted as discrete squares in Figure 3 which also shows the elements of the rotated right hand side vector $\mathbf{U}^T\mathbf{b}$ plotted as circles connected by straight line segments. The decomposition was computed by subroutine SGESVD from the LAPACK collection [1], using double precision arithmetic on a Sparc-20 workstation. The largest and smallest singular values are

$$\sigma_1(\mathbf{A}) = 2.882 \times 10^5 \; , \quad \sigma_{241}(\mathbf{A}) = 2.622 \times 10^{-2} \; , \tag{6.8}$$

which gives $\mathrm{cond}(\mathbf{A}) = 1.099 \times 10^7$. The relative acccracy of the arithmetic was $\epsilon_{\mathrm{mach}} = 2.22 \times 10^{-16}$, so the smallest singular value is many orders of

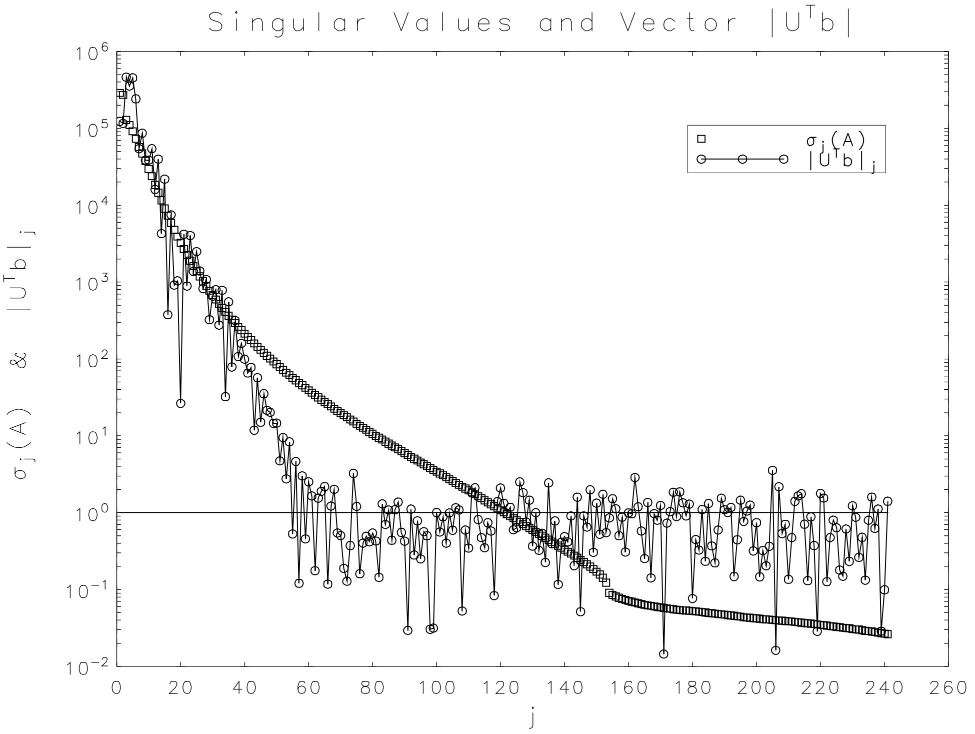

Figure 3: Singular values (squares) and first n elements of $|\mathbf{U}^T\mathbf{b}|$ (circles) for the modified Phillips problem

magnitudes larger than the effective zero level. Thus there is no reason to assume that rank$(\mathbf{A}) < n$. Yet the problem obviously needs some sort of truncation to prevent the estimated solution from capturing variance that properly belongs in the residuals. Fortunately, the desired result can be obtained by leaving the singular values unchanged and truncating the rotated right hand side vector $\mathbf{U}^T\mathbf{b}$.

7 Truncating the Vector $\mathbf{U}^T\mathbf{b}$

Even a cursory inspection of the $|\mathbf{U}^T\mathbf{b}|_i$ plotted in Figure 3 reveals a sharp dichotomy in the distribution at $i = 55$. Before the break, the upper envelope of the distribution is decreasing at a faster rate than that of the singular values. After the break, the distribution is quite flat with 64 of the 186 values above the horizontal line $|\mathbf{U}^T\mathbf{b}| = 1$ and 122 values below. That line is, by Equations (5.3), just the one standard deviation level for the random errors $(\mathbf{U}^T\boldsymbol{\eta})_i$. If the final 186 elements of $\mathbf{U}^T\mathbf{b}$ were chosen randomly from an $n(0,1)$ distribution,

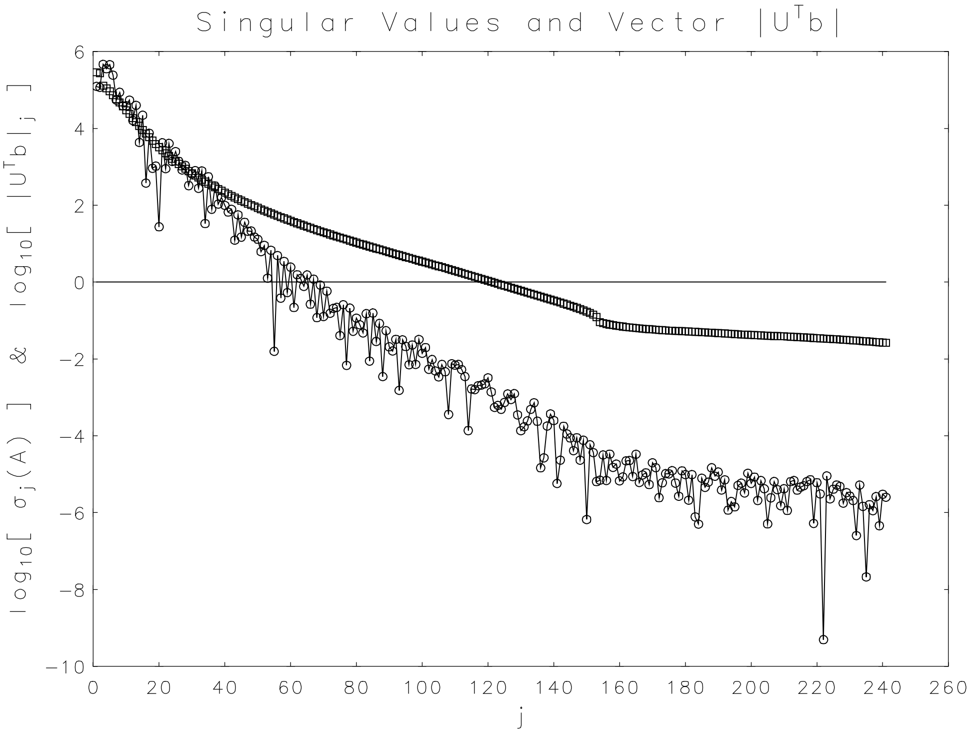

Figure 4: Singular values (squares) and first n elements of $|\mathbf{U}^T\mathbf{b}|$ (circles) for the modified Phillips problem with no random errors

then one would expect roughly 62 of them to satisfy $|\mathbf{U}^T\mathbf{b}|_i > 1$ and the other 124 to satisfy $|\mathbf{U}^T\mathbf{b}|_i \leq 1$. Since that is almost exactly what is observed, it is quite reasonable to conclude that the flat part of the distribution represents elements of $\mathbf{U}^T\mathbf{b}$ that are dominated by the measuring errors. This conclusion is reinforced by Figure 4 which gives plots of the $\sigma_i(\mathbf{A})$ and $|\mathbf{U}^T\mathbf{b}|_i$ for the same problem with no random errors in the $\mathbf{b}$ vector. The singular values are identical to the ones in Figure 3, but the distribution of the $|\mathbf{U}^T\mathbf{b}|_i$ satisfies the discrete Picard condition over the whole range, and Eq. (5.9) yields estimates of the elements of $\mathbf{x}^*$ that are correct to 8 significant digits.

The foregoing observations suggest a more logical way to truncate the singular value decomposition. *Rather than zeroing some of the singular values one should instead zero those elements of* $|\mathbf{U}^T\mathbf{b}|$ *that are judged to be mostly random error.* The idea is to pick a truncation level τ and require the estimated solution $\tilde{\mathbf{x}}$ to satisfy

$$(\mathbf{V}^T\tilde{\mathbf{x}})_i = \begin{cases} \dfrac{\left(\mathbf{U}^T\mathbf{b}\right)_i}{\sigma_i} & , \quad \text{if } |\mathbf{U}^T\mathbf{b}|_i > \tau \ , \\ 0 & , \quad \text{otherwise} \ , \end{cases} \qquad i = 1, 2, \ldots, n \ . \tag{7.1}$$

The sum of squared residuals is then given by

$$r^2 = ||\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}||^2 = \sum_{i \in \mathcal{I}} \left(\mathbf{U}^T\mathbf{b}\right)_i^2 \ , \tag{7.2}$$

where the indexing set for the sum is

$$\mathcal{I} = \left\{ \ i \ \middle| \ |\mathbf{U}^T\mathbf{b}|_i \le \tau \ , \ i = 1, 2, \ldots, n \ \right\} \bigcup \{ \ n+1, \ n+2, \ \ldots, \ m \ \} \ . \tag{7.3}$$

If $\mathbf{u}_i$ is the ith column of $\mathbf{U}$, then Eq. (7.1) can be written

$$(\mathbf{V}^T\tilde{\mathbf{x}})_i = \begin{cases} \dfrac{\left(\mathbf{u}_i^T\mathbf{b}\right)}{\sigma_i} & , \quad \text{if } |\mathbf{u}_i^T\mathbf{b}| > \tau \ , \\ 0 & , \quad \text{otherwise} \ , \end{cases} \qquad i = 1, 2, \ldots, n \ , \tag{7.4}$$

and if an $m \times n$ matrix $\tilde{\mathbf{U}}_1$ is defined by

$$\tilde{\mathbf{U}}_1 = (\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \cdots, \tilde{\mathbf{u}}_n) \ , \quad \text{where} \qquad \tilde{\mathbf{u}}_i = \begin{cases} \mathbf{u}_i & , \quad \text{if } |\mathbf{u}_i^T\mathbf{b}| > \tau \ , \\ \mathbf{0} & , \quad \text{otherwise} \ , \end{cases} \tag{7.5}$$

then the estimated solution and residual vectors can be written

$$\tilde{\mathbf{x}} = \mathbf{V}\mathbf{\Sigma}^{-1}\tilde{\mathbf{U}}_1^T\mathbf{b} \ , \quad \text{and} \qquad \tilde{\mathbf{r}} = \left(\mathbf{I}_m - \tilde{\mathbf{U}}_1\tilde{\mathbf{U}}_1^T\right)\mathbf{b} \ . \tag{7.6}$$

It follows from a fundamental theorem in multivariate analysis [2, Thm. 2.4.5] that

$$\tilde{\mathbf{r}} \sim N\left[\left(\mathbf{I}_m - \tilde{\mathbf{U}}_1\tilde{\mathbf{U}}_1^T\right)\mathbf{A}\mathbf{x}^* \ , \ \ \mathbf{I}_m - \tilde{\mathbf{U}}_1\tilde{\mathbf{U}}_1^T \ \right] \ . \tag{7.7}$$

It is clear that as τ increases, the number of non-zero columns in $\tilde{\mathbf{U}}$ decreases, so the variance matrix for $\tilde{\mathbf{r}}$ more nearly approximates a diagonal matrix.

The success of the proposed method depends crucially on the choice of the truncation level τ. A safe and effective approach is to try several values of τ and use the diagnostics described in Section 4 and Appendix D to make the final choice. The results of this strategy for the modified Phillips problem are summarized in Table 1. The labels for the first, second and final columns are

Table 1: Estimate diagnostics for the modified Phillips Problem

τ	$\sum(\mathbf{b}-\mathbf{A}\tilde{\mathbf{x}})_i^2$	$\%(\mathcal{C}_k \in \mathcal{B}_{95})$	Length$\{\mathcal{C}\}$	$\lvert\Delta x\rvert_{rms}$	Figure
0.0	43.0	31.6	1.238	14.937	2
3.0	235.6	79.5	1.211	5.6865	5
4.0	267.7	100.0	1.205	0.0106	6
4.7	289.2	100.0	1.209	0.0122	
5.0	311.4	100.0	1.207	0.0132	
9.0	380.7	100.0	1.206	0.0148	

self explanatory. The third column contains the percentage of the 4096 cumulative periodogram ordinates that lie inside the 95% confidence band for white noise. The fourth column contains the length of the cumulative periodogram plot which can be compared with the length 1.11803 for a pure white noise series. The fifth column gives the root mean square average magnitude of the errors in the elements of the estimated solution vector, i.e.,

$$|\Delta x|_{rms} = \sqrt{\frac{1}{n}\sum_{j=1}^{n} |\tilde{x}_j - x_j^*|^2}\ . \tag{7.8}$$

It can be included only because the true solution is known. The first row in the table gives the results for the least squares estimate shown in Figure 2.

A guideline for choosing a lower bound for the value of τ is the fact that most experimentalists would be reluctant to claim that a measured value is significantly different from zero if its magnitude does not exceed 3 standard deviations for the error in the measurement. Since each $|\mathbf{U}^T\mathbf{b}|_i$ is scaled in units of one standard deviation of its own random error, it suffices to choose $\tau = 3.0$. The results obtained for this choice are given in Figure 5 and row 2 of the table. The sum of squared residuals, 235.6, is less than the 1 percentile value for the $\chi^2(300)$ distribution (see column 6 of Table 3, in Appendix A), i.e.,

$$\Pr\left\{\|\mathbf{b}-\mathbf{A}\mathbf{x}^*\|^2 \le 235.6\right\} < \Pr\left\{\|\mathbf{b}-\mathbf{A}\mathbf{x}^*\|^2 \le 245.3\right\} = 0.01\ . \tag{7.9}$$

Furthermore, only 79.5% of the cumulative periodogram ordinates lie inside the 95% band for white noise. The plot of the estimate reveals the persistence of strong high frequency oscillations which, by inspection of Figure 3, can be attributed to the component $|\mathbf{U}^T\boldsymbol{b}|_{205} = 3.547$. Repeatedly measuring a standard normally distributed quantity will give a measurement with magnitude

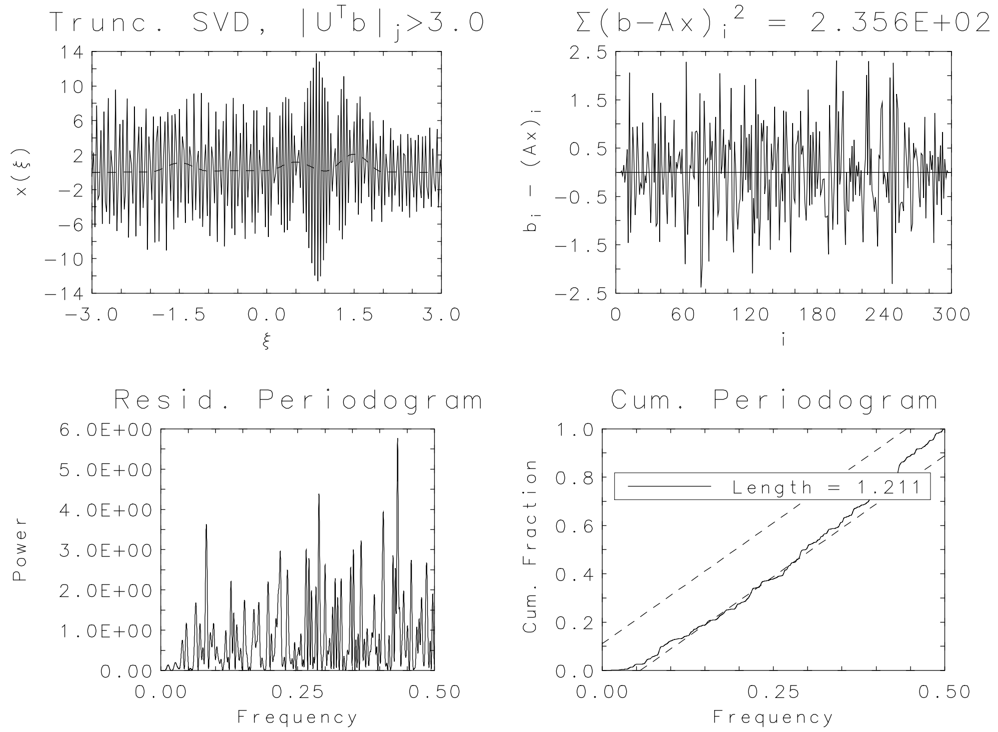

Figure 5: Truncated solution and diagnostic plots for the modified Phillips problem with $\tau = 3$

greater than 3 roughly once in every 385 measurements so the occurrence of one such value in the final 186 $|\mathbf{U}^T\boldsymbol{b}|_i$ is not an unlikely event. Its effect can be transferred from the estimated solution to the residuals by increasing the truncation threshold to $\tau = 4$. The results for that value are given in Figure 6 and row 3 of the table. The sum of squared residuals, 267.7, differs from the expected value by about 1.3 standard deviations. While this is not a small difference, it is clear from column 6 of Table 3 that

$$0.05 \;=\; \Pr\left\{\|\mathbf{b}-\mathbf{A}\mathbf{x}^*\|^2 \leq 260.6\right\} < \;\Pr\left\{\|\mathbf{b}-\mathbf{A}\mathbf{x}^*\|^2 \leq 267.7\right\} \;, \qquad (7.10)$$

so 267.7 must be regarded as an acceptable value. This conclusion is reinforced by the fact that none of the cumulative periodogram ordinates lie outside the 95% band. The plot of the estimated solution, at the upper left of Fig. 6, is so close to that of the true solution that the two are almost indistinguishable.

Since there are only 241 discrete $|\mathbf{U}^T\mathbf{b}|_i$, there are only 241 possible truncated estimates $\tilde{\mathbf{x}}$. Only 55 of them correspond to truncation levels greater than $\tau = 4$. The smallest four of those are

$$|\mathbf{U}^T\mathbf{b}|_{56} = 4.63, \;\; |\mathbf{U}^T\mathbf{b}|_{51} = 4.71, \;\; |\mathbf{U}^T\mathbf{b}|_{54} = 8.33, \;\; |\mathbf{U}^T\mathbf{b}|_{52} = 9.49. \quad (7.11)$$

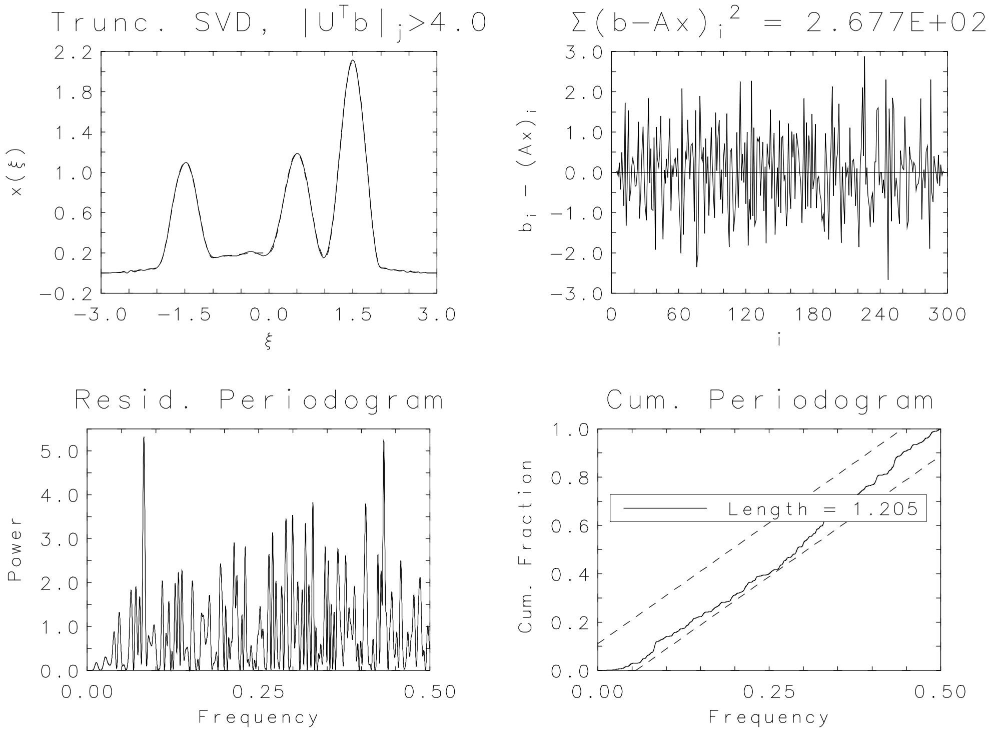

Figure 6: Truncated solution and diagnostic plots for the modified Phillips problem with $\tau = 4$

These values define 3 ranges for τ, in each of which the estimates are all identical. For the lowest 2 ranges, the results for $\tau = 4.7$ and $\tau = 5.0$ are given in the fourth and fifth rows of Table 1. In both cases, the estimated sum of squared residuals lie within one standard deviation of the expected value, and all of the cumulative periodogram ordinates lie in the white noise band. The highest range is represented by $\tau = 9$ with the results summarized in the row 6 of the table. Although all of the periodogram ordinates lie within the 95% confidence band for white noise, the sum of squared residuals, 380.7, is much too large to be acceptable since, from column 6 of Table 3,

$$\begin{aligned} \Pr\{\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2 > 380\} &< \Pr\{\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2 > 359.1\} \\ &= 1 - \Pr\{\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2 \leq 359.1\} \\ &= 0.01\ . \end{aligned} \tag{7.12}$$

Thus, the only estimates to give acceptable values for the sum of squared residuals are the ones corresponding to truncation levels $\tau = 4.0,\ 4.7,$ and 5.0. Since the one for $\tau = 4.0$ gave the smallest Length$\{\mathcal{C}\}$, it was chosen as the most acceptable, and this choice is validated by the fact that it also gave the smallest $|\Delta x|_{rms}$. But the other two estimates are quite good also, and are qraphically almost indistinguishable from the $\tau = 4.0$ estimate.

8 A Real-World Example

The measurement of nuclear radiation spectra is one important source of first kind integral equations with uncertainties. Consider the energy spectrum of the neutrons produced by the nuclear reaction

$$T \ (d \ , \ n \ ) \ ^4He \ , \tag{8.1}$$

i.e. tritium nuclei bombarded with deuterons to produce helium nuclei and the neutrons being measured. If the bombarding particles are monoenergetic then so are the neutrons produced, with energy in the range 12 - 22 MeV (million electron volts), depending on that of the bombarding particles. Although the neutrons are monoenergetic, the measuring instrument both smears and distorts the expected spectrum. The instrument in question, an NE-213 spectrometer, has been described by Verbinski, et. al. [20] and Burrus and Verbinski [5]. Its response functions are plotted in Figure 7, where incident energy corresponds to the variable ξ and pulse height to the variable t in Eqs. (1.1). The incident neutrons are absorbed by a plastic scintillator which emits a light pulse whose size is proportional to the incident energy. The light pulse is in turn detected and amplified by a photomultiplier tube which transmits a corresponding electrical pulse to an electronic circuit which then increments a counter in one of $m = 133$ pulse-height bins designed to cover and subdivide the energy range of all possible incident neutrons. The integral equations modelling the process are

$$c_i \ = \int_0^{E_{up}} K_i(E)N(E)dE \ \ + \ \ \varepsilon_i \ , \quad i = 1, 2, \ldots, 113 \ , \tag{8.2}$$

where c_i is the number of pulses counted in the ith bin, $N(E)$ is the unknown number of neutrons at energy E, and the $K_i(E)$ are the instrument response functions for the whole detector system. For a given value of E, the quantity $K_i(E)dE$ is the probability that a neutron with energy in the range $E \pm \frac{1}{2}dE$ would produce a light pulse that would increment the count in the ith pulse-height bin. Hopefully, the bin number in which a neutron is counted would increase as the energy increases, but an inspection of Figure 7 shows that this does not always happen. Ideally the $K_i(E)$ should be a series of narrow peaks distributed along a diagonal line running from the upper left to the lower right of the energy, pulse-height domain. The figure does exhibit a ridge along that

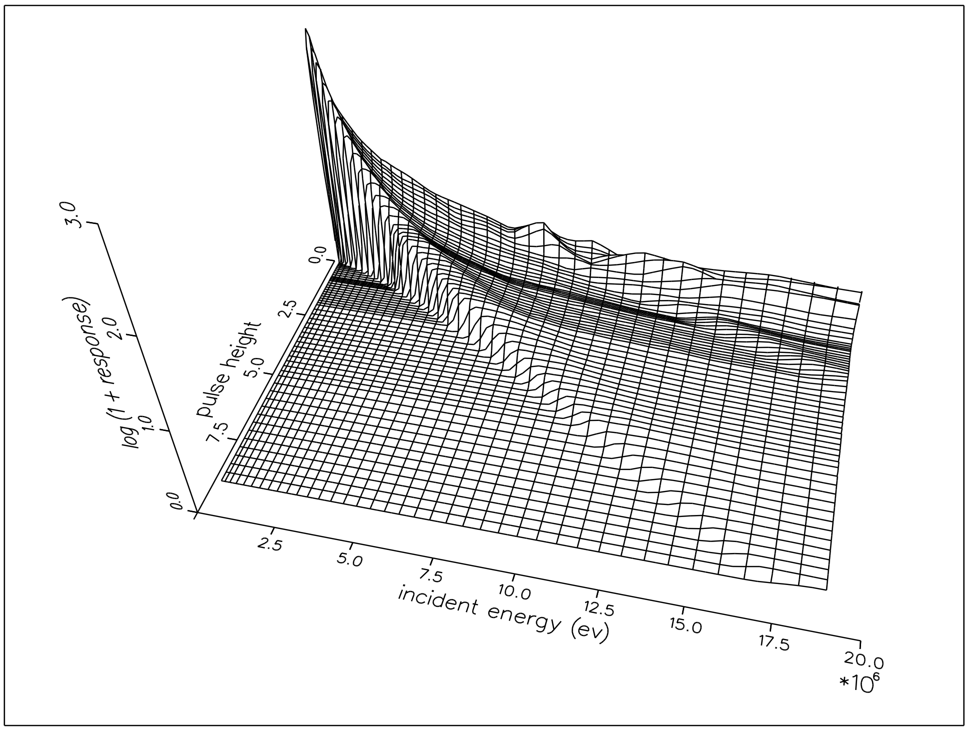

Figure 7: Instrument response functions for the Burrus neutron spectrum problem. The quantity $\log_{10}[1 + K_i(E)]$ is plotted in order to more clearly show the structure for higher energies.

direction, but it attenuates to a barely discernible ripple for higher energies. Even worse, the long energy tails for the lower pulse-height bins make it more likely that a higher energy neutron will produce an increment in a lower rather than a higher bin.

Figure 8 shows a measured spectrum obtained for the neutrons from the reaction (8.1). Note that count rate rather than counts are plotted against pulse-height. The experimental procedure is to accumulate counts for a time sufficient to allow good estimates to be made for the uncertainties in the count totals and then to divide by the elapsed time to get rates. In Eq. (8.2) this division replaces the c_i with count rates y_i and the number $N(E)$ of neutron

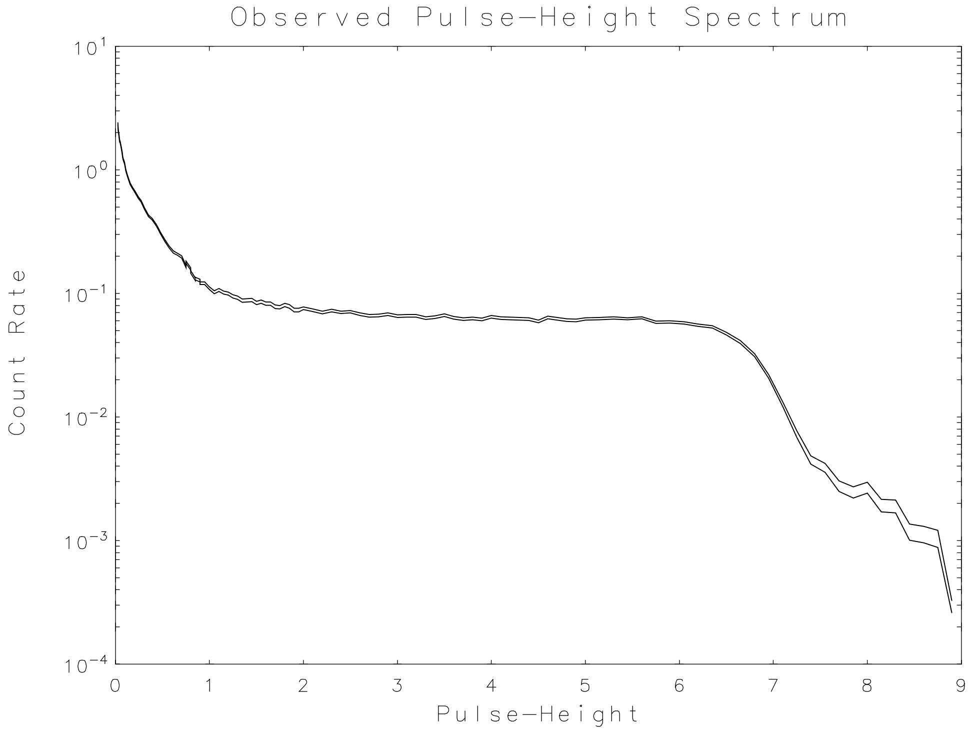

Figure 8: Measured pulse-height spectrum for the Burrus problem. The two piecewise linear curves are the $\pm$ one standard deviation bounds for the measured values

with a neutron flux $x(E)$. The result is

$$y_i \;=\int_0^{E_{up}} K_i(E)x(E)dE \;\;+\;\; \epsilon_i \;, \quad i = 1,2,\ldots,113\;, \tag{8.3}$$

where the ϵ_i are noise rates obtained from the ε_i. Estimates for the standard deviations of the latter quantities were obtained from the usual assumption about counting errors which is that the standard deviation for any bin is approximately equal to the square root of the number of counts in that bin. Although this procedure assumes a Poisson distribution for the error in any bin, the number of counts accumulated was sufficient to permit the use of the normal approximation. Assuming that the measuring errors in any two bins are statistically independent and dividing the estimated standard deviations by the time of accumulation gives a diagonal variance matrix $\mathbf{S}^2$ for the errors.

When the energy E is equated to the variable ξ, the equations (8.3) become identical to those in (1.2). They were discretized using simple rectangular quadrature with an $n = 77$ point energy mesh, $0.2\ MeV = E_1 < E_2 < \cdots <$

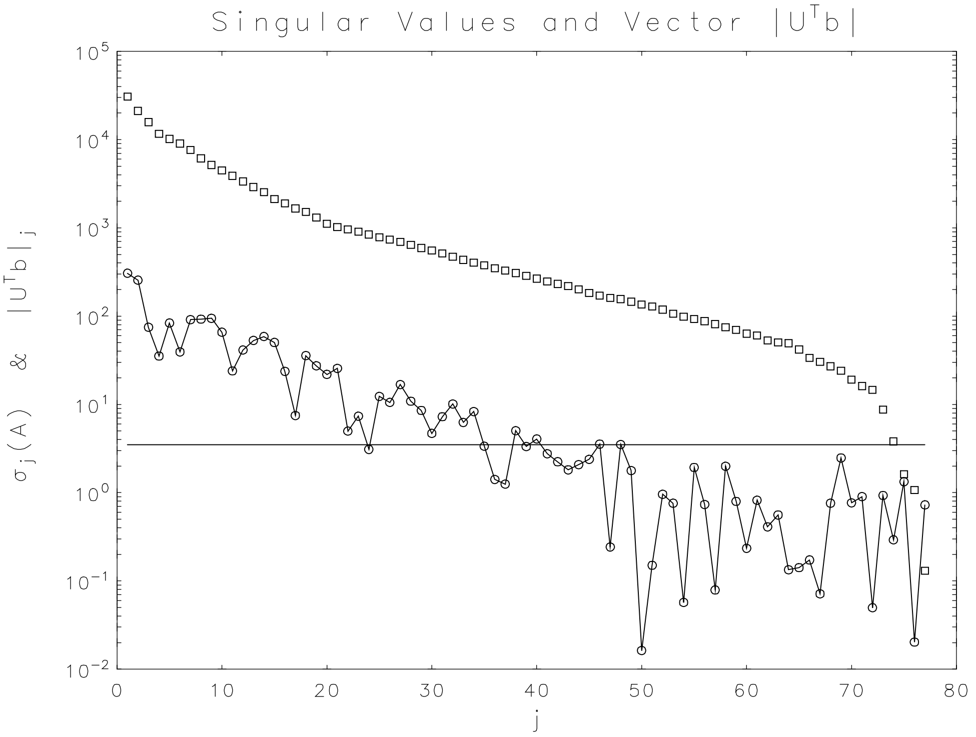

Figure 9: Singular values (squares) and first n elements of $|\mathbf{U}^T\mathbf{b}|$ (circles) for the Burrus neutron spectrum problem. The horizontal line represents the truncation level $\tau = 3.5$.

$E_{77} = 18.91\ MeV$, to give the linear regression model (1.3). Multiplying through by $\mathbf{S}^{-1}$ gave the normalized model (2.3). Although the two curves plotted in Figure 8 indicate good statistics for the measurements, there is no sign of the dominant peak expected for monoenergetic neutrons. The sharp rise in the lower pulse-height bins contains most of the counts that should have gone into that missing peak. The least squares estimate, which is overwhelmed by the amplified noise, gives

$$r_{\min}^2 = \|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|^2 = 38.05 \tag{8.4}$$

which is almost 5 standard deviations smaller than the expected value 113. The singular values and first n elements of $|\mathbf{U}^T\mathbf{b}|$ are shown in Figure 9. The condition number of the matrix is 2.3503×10^5 so it is clearly not rank-deficient. The distribution of the $|\mathbf{U}^T\mathbf{b}|_i$ exhibits a dichotomy at $i = 47$ where it first drops below the one sigma level for the error. Table 2 gives estimate diagnostics for several values of τ. The quantities tabulated are the same as those in the first 4 columns of Table 1 All of the non-zero τ give estimates

Table 2: Estimate diagnostics for the Burrus neutron spectrum problem

τ	$\sum(\mathbf{b}-\mathbf{A}\mathbf{x})_i^2$	$\%(\mathcal{C}_k \in \mathcal{B}_{95})$	Length$\{\mathcal{C}(f)\}$
0.0	38.0	37.4	1.255
3.0	93.8	100.0	1.211
3.5	125.7	100.0	1.212
4.0	150.7	100.0	1.211
5.0	213.9	63.6	1.216

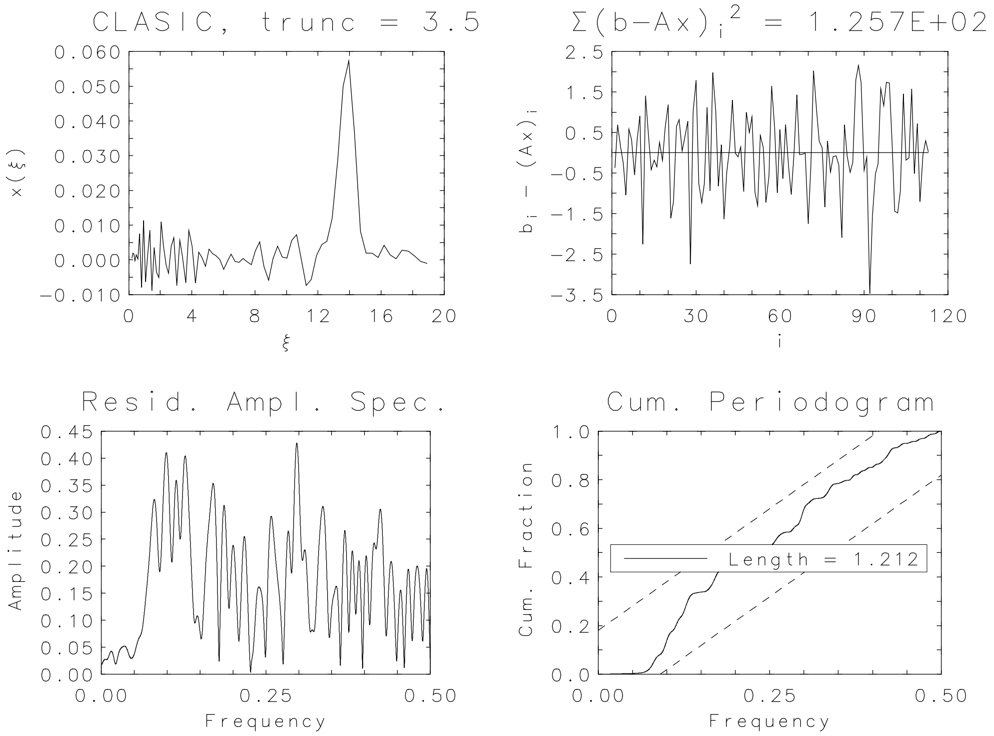

Figure 10: Truncated solution and diagnostic plots for the Burrus neutron spectrum problem with $\tau = 3.5$

with a strong peak at approximately 14 MeV, but only $\tau = 3.0,\ 3.5,$ and 4.0 give residuals consistent with white noise. The sum of squared residuals for $\tau = 3.5$ is closest to the expected value, so it was chosen as optimal even though the length of the cumulative periodogram was slightly larger than those for $\tau = 3.0$ and 4.0. The estimate and diagnostic plots for $\tau = 3.5$ are given in Figure 10. The estimate is dominated by a single peak centered at 14 MeV which is just what was expected for monoenergetic neutrons.

9 Discussion and Conclusions

The primary message of this paper is that for ill-posed problems in which the dominant errors are the measurement uncertainties for the right hand side vector, the instability in estimating the solution is attributable to components of the measurements which are overwhelmed by the random error. In general, these components will not correspond exactly to specific elements of the right hand side vector. Scaling the problem by $\mathbf{S}^{-1}$ reduces it to one in which the errors all have unit variance, and premultiplying the scaled measurement vector by $\mathbf{U}^T$ rotates it into a basis in which the noise dominated components can be isolated by well understood statistical criteria. Zeroing the noise dominated components stabilizes the solution estimate without altering the kernel matrix which is assumed to be known more accurately than the right hand side vector. Of course this procedure does not guarantee that the estimated solution will be close to the true solution since the noise may be large enough to overwhelm important components of the signal. The only way to retrieve such components would be to repeat the measurments with more accuracy.

The traditional method of truncating the singular value decomposition by zeroing some of the singular values amounts to altering a well known quantity, $\mathbf{A}$, in order to compensate for uncertainties in an unknown quantity $\mathbf{b}^* = \mathbf{A}\mathbf{x}^*$. In the process it reassigns the rank of the matrix, and since the truncation level depends crucially on the size of the errors in the right hand side vector, it follows that the altered rank depends on the size of those errors. As an example, consider Figure 11 which is a plot of the σ_i and $|\mathbf{U}^T\mathbf{b}|_i$ for a problem obtained from the one illustrated in Figure 3 by increasing all the errors by a factor of 10. Both figures display the flat segment of the $|\mathbf{U}^T\mathbf{b}|$ curve, which corresponds to error saturation, but for the problem with the larger errors, this flat segment intersects the singular value distribution at a point higher and to the left of that for the problem with the smaller errors. Thus, if the truncation required to stabilize the estimate were accomplished by zeroing singular values, the problem with larger errors would assign a smaller numerical

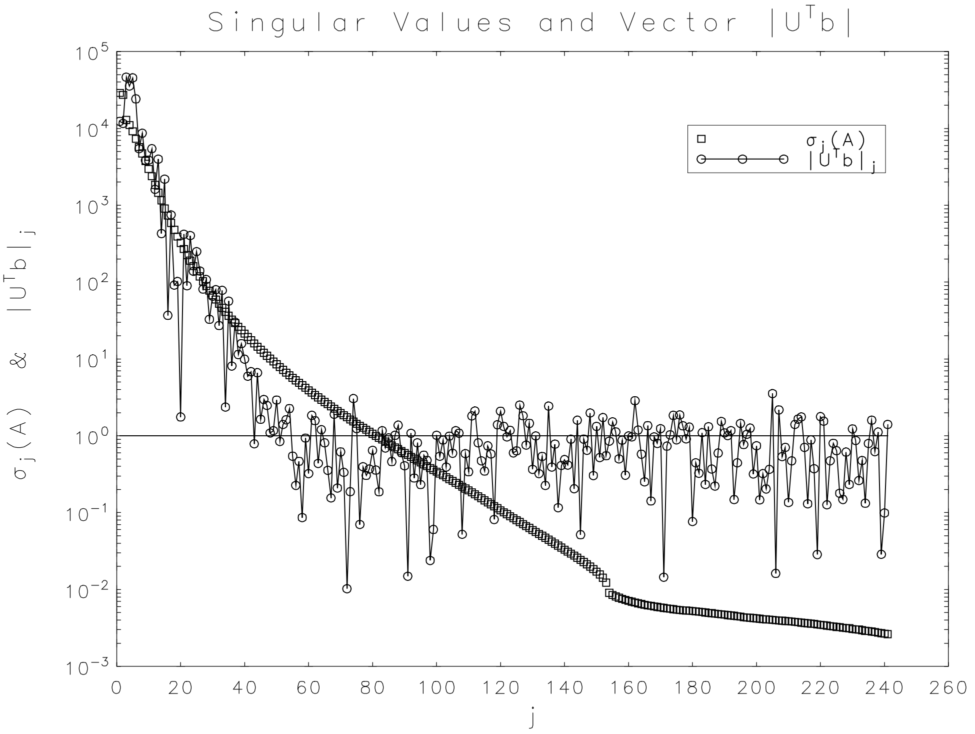

Figure 11: Singular values (squares) and first n elements of $|\mathbf{U}^T\mathbf{b}|$ (circles) obtained by increasing the errors in the modified Phillips problem by a factor of 10

rank to the matrix than the one with smaller errors. Although it seems absurd to have the rank of a matrix depend on the right hand side vector, precisely this suggestion was made quite recently by P.C. Hansen [12, Chapt. 7] who defined an "effective numerical rank" which is not even an integer. A better approach might be to simply abandon the notion of "numerical rank", but the huge vested interest in rank determination and rank revealing decompositions will probably preclude such a course of action.

The truncation procedure described in this paper depends critically on the availability of good estimates for the uncertainties in the elements of the measured vector. A good experimenter will usually provide such estimates, but it is not uncommon for them to be systematically too small or, more rarely, too large. Often the form of the functional dependence of the errors on the magnitudes of the measurements is correct but the estimated standard deviations are all too small, or too large, by a constant factor. In such cases, the diagnostics introduced in Section 4 can be used to pinpoint the discrepancy and rectify it. The usual symptom is that the threshold τ at which the estimate begins to be stabilized is quite different from 3 or 4 and also different from the value which gives residuals closest to white noise, as measured by the length of the cumulative periodogram. In such cases it is a straightforward though often tedious matter to adjust the scale of the estimated errors by a constant factor which brings the two values of τ into agreement somewhere in the interval $3 \leq \tau \leq 5$.

The situation is more difficult when no estimates are given for the measurement errors. The usual approach is to implicitly assume that $\mathbf{S}^2 = s^2\mathbf{I}_m$ where s is an unknown constant which can be estimated from the least squares sum of squared residuals. Of course the latter quantity is usually much smaller than the expected value m, but the size of s can then be adjusted upward to a value such that some truncation level in the range $3.0 \leq \tau \leq 5.0$ gives residuals close to white noise and a sum of squared residuals close to the expected value. This usually iterative technique works well if the standard deviations of the measured errors are truly constant. If not, it may still be possible to assume a functional form for the uncertainties which will determine the matrix $\mathbf{S}^2$ up to an unknown scaling constant which can then be adjusted for consistency in the τ values. The disadvantage in this approach is the amount of work involved in validating a functional form and adjusting the scaling, but sometimes the results obtained justify the effort.

A number of people have suggested that the truncation procedure advocated in this paper is incomplete because it has not been accompanied by an algorithm for automatically determining the optimum τ, using some criterion like those used in choosing the smoothing parameter for regularization methods, e.g., generalized cross-validation, or the L-curve method. In practice, for any given problem, only a few discrete truncations give estimates with residuals that have both a reasonable sum of squares and an acceptable spectral distribution, and it is often the case that the two criteria are optimized by different truncation levels. Thus some judgement is required to choose which truncation level to use. Also there is little to be gained computationally by confining the calculations to a single τ. The amount of work required is dominated by the computation of the singular value decomposition, but once that is accomplished it is relatively easy to isolate the few possibly acceptable truncation levels and, using a fast Fourier transform algorithm, to calculate all of the diagnostics. Since the total effort involved is small in comparison to that required to make the measurements, most users with real world problems will not be unduly troubled by having to make the final choice of the optimum value. Of course the procedure can be automated by calculating the solutions and diagnostics on a suitably chosen mesh for τ. If the mesh spacing is fine enough to find all of the possibly acceptable solutions, it will produce many duplicate solutions, but these will not seriously complicate the final choice of the parameter value. Users of regularization methods might also be well advised to consider the same two residual diagnostics in evaluating their favorite criteria for choosing the regularization parameter.

Acknowledgements

I would like to thank Drs. W.R. Burrus and R.E. Funderlic for teaching me the pitfalls of truncating the singular value distribution, Drs. D.P. O'Leary and K.A. Remington for their reviews and suggestions for improving the numerical analysis, and Dr. M.G. Vangel for many useful suggestions for improving the statistical aspects of this manuscript.

Appendices

A Percentiles for the $\chi^2(m)$ Distribution

The cumulative distribution funtion for $\chi^2(m)$ is defined by

$$F(\chi^2) = \int_0^{\chi^2} \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} z^{\frac{m-2}{2}} \exp\left(-\frac{z}{2}\right) dz \ , \tag{A.1}$$

and for any probability α, the α-point of the distribution can be found by solving $F(\chi^2) = \alpha$. If the solution is χ^2_α, then

$$\Pr\left\{\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2 \leq \chi^2_\alpha\right\} = \alpha \ , \tag{A.2}$$

and

$$\Pr\left\{\|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|^2 > \chi^2_\alpha\right\} = 1 - \alpha \ . \tag{A.3}$$

Standard tables (e.g., [3]) of the χ^2-distribution list χ^2_α for various values of α and m, up to about $m = 30$. For $m > 30$, the quantity $\left(\sqrt{2\chi^2} - \sqrt{2m-1}\right)$ is distributed approximately like the standard normal distribution, so χ^2_α is well approximated by

$$\chi^2_\alpha = \frac{1}{2}\left[\kappa_\alpha + \sqrt{2m-1}\right]^2 \ , \tag{A.4}$$

where κ_α is the α-point of the standard normal distribution, i.e.,

$$\int_{-\infty}^{\kappa_\alpha} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz = \alpha \ . \tag{A.5}$$

Table 3: Selected percentiles for the standard normal and χ^2 distributions

α	κ_α	χ^2_α			
		$m=100$	$m=150$	$m=200$	$m=300$
0.01	-2.326	69.4	112.0	155.7	245.3
0.05	-1.645	77.7	122.4	168.0	260.6
0.33	-0.440	93.4	142.0	190.8	288.8
0.50	0.0	99.5	149.5	199.5	299.5
0.67	0.440	105.8	157.2	208.4	310.4
0.95	1.645	124.1	179.3	233.7	341.1
0.99	2.326	135.0	192.4	248.7	359.1

Table 3 gives, for various values of α, the corresponding values of κ_α and, for some representative values of m, the estimates of χ^2_α calculated by Eq. (A.4). Good public domain software for computing χ^2_α can be found in the GAMS collection [7, Class L5a1c].

B A Variant of the Phillips Problem

A useful test problem which shares many of the characteristics of real instrument correction problems is obtained by discretizing a variant of the well known [16] Phillips equation. This modified Phillips problem can be written

$$y(t) = \int_{-3}^{3} K(t,\xi)x(\xi)\,d\xi\ , \quad -6 \le t \le 6\ , \tag{B.1}$$

with the kernel given by

$$K(t,\xi) = \begin{cases} \frac{1}{6}\left\{1+\cos\left[\frac{\pi}{3}(\xi - t)\right]\right\} & , \quad |\xi - t| \le 3\ , \quad |t| \le 6\ , \\ 0 & , \quad \text{otherwise}\ , \end{cases} \tag{B.2}$$

and the exact solution by

$$x(\xi) = \beta(\xi) + \sum_{k=1}^{3} c_k(\xi)\ , \tag{B.3}$$

where

$$\beta(\xi) = \begin{cases} A_0\left[1+\cos\left(\frac{\pi}{3}\xi\right)\right] & , \quad |\xi| \le 3\ , \\ 0 & , \quad |\xi| > 3\ , \end{cases} \tag{B.4}$$

and

$$c_k(\xi) = \begin{cases} A_k \left\{1 + \cos\left[2\pi\left(\xi - \psi_k\right)\right]\right\} & , \quad |\xi - \psi_k| \le \frac{1}{2} \ , \\ 0 & , \quad \text{otherwise} \ , \end{cases} \tag{B.5}$$

with amplitude constants A_k and centering constants ψ_k chosen to be

$$\begin{array}{llll} A_0 = 0.1 \ , & A_1 = 0.5 \ , & A_2 = 0.5 \ , & A_3 = 1.0 \ , \\ & \psi_1 = -1.5 \ , & \psi_2 = 0.5 \ , & \psi_3 = 1.5 \ . \end{array} \tag{B.6}$$

The kernel differs from the Phillips original only in the inclusion of the normalizing factor $\frac{1}{6}$ which was added to assure that for any ξ,

$$\int_{-3+\xi}^{3+\xi} K(t,\xi)\, dt \ = \ 1 \ . \tag{B.7}$$

For a measuring instrument this condition assures that conservation laws are not violated. Plots of $K(t,\xi)$ for 5 representative values of ξ are given on the left in Figure 1.

The exact solution to the original Phillips problem appears in a scaled down form as the $\beta(\xi)$ term in the solution to the new problem. The scaling constant A_0 was chosen to reduce the original solution to the role of a background function on which to superimpose the three discrete peaks represented by the $c_k(\xi)$ terms. The three points $\xi = \psi_k$ are the centers of these peaks and the constants $2A_k$ are their heights above the background. The new solution function is plotted as a dashed line on the right in Figure 1.

These changes in the Phillips problem were designed to make it more challenging and more reminiscent of real-world instrument correction problems. Unfortunately they also make the representation of the function $y(t)$ more complicated. Substituting (B.3) into (B.1) gives

$$y(t) = \int_{-3}^{3} K(t,\xi)\beta(\xi)d\xi + \sum_{k=1}^{3} \int_{-3}^{3} K(t,\xi)c_k(\xi)d\xi \ , \ \ -6 \le t \le 6 \ , \tag{B.8}$$

but care must be exercised in evaluating these integrals because $K(t,\xi) = 0$ on half of the rectangular domain $\{\ (t,\xi)\ |-6 \le t \le 6 \ , \ -3 \le \xi \le 3\ \}$ and each of the $c_k(\xi)$ is zero everywhere except on the interval $\psi_k - \frac{1}{2} \le \xi \le \psi_k + \frac{1}{2}$. The last equation can also be written

$$y(t) = B(t) + \sum_{k=1}^{3} C_k(t) \ , \tag{B.9}$$

where

$$B(t) \equiv \int_{-3}^{3} K(t,\xi)\beta(\xi)d\xi \ , \tag{B.10}$$

and

$$C_k(t) \equiv \int_{\psi_k-\frac{1}{2}}^{\psi_k+\frac{1}{2}} K(t,\xi)c_k(\xi)d\xi \ , \quad k=1,\ 2,\ 3\ . \tag{B.11}$$

Evaluating the integral for $B(t)$ gives

$$B(t) = \frac{1}{6}A_0\left\{(6-|t|)\left[1+\frac{1}{2}\cos\left(\frac{\pi}{3}t\right)\right]+\frac{9}{2\pi}\sin\left(\frac{\pi}{3}|t|\right)\right\}\ , \quad -6\le t\le 6\ , \tag{B.12}$$

and the integrals for $C_k(t)$ can be written

$$C_k(t) = \frac{1}{6}A_kL_k(t)\ , \quad k=1,\ 2,\ 3\ , \tag{B.13}$$

where

$$L_1(t) = \begin{cases} 0\ , & -6\le t\le -5, \\[2ex] t+5+\frac{1}{2\pi}\sin\left[\pi(2t+9)\right]+\frac{3}{\pi}\sin\left[\frac{\pi}{3}(t+2)\right] & \\ \quad +\frac{3}{10\pi}\left\{\sin\left[\pi(2t+8)\right]-\sin\left[\frac{\pi}{3}(t-1)\right]\right\} & \\ \quad +\frac{3}{14\pi}\left\{\sin\left[\pi(2t+10)\right]+\sin\left[\frac{\pi}{3}(t+5)\right]\right\}, & -5\le t\le -4, \\[2ex] 1+\frac{3}{\pi}\left\{-\sin\left[\frac{\pi}{3}(1+t)\right]+\sin\left[\frac{\pi}{3}(2+t)\right]\right\} & \\ \quad +\frac{3}{10\pi}\left\{\sin\left[\frac{\pi}{3}(4+t)\right]+\sin\left[\frac{\pi}{3}(1-t)\right]\right\} & \\ \quad +\frac{3}{14\pi}\left\{\sin\left[\frac{\pi}{3}(2-t)\right]+\sin\left[\frac{\pi}{3}(5+t)\right]\right\}, & -4\le t\le 1, \\[2ex] 2-t+\frac{1}{2\pi}\sin\left[\pi(3-2t)\right]-\frac{3}{\pi}\sin\left[\frac{\pi}{3}(t+1)\right] & \\ \quad +\frac{3}{10\pi}\left\{\sin\left[\frac{\pi}{3}(t+4)\right]+\sin\left[\pi(2-2t)\right]\right\} & \\ \quad +\frac{3}{14\pi}\left\{\sin\left[\frac{\pi}{3}(2-t)\right]+\sin\left[\pi(4-2t)\right]\right\}, & 1\le t\le 2, \\[2ex] 0\ , & 2\le t\le 6, \end{cases} \tag{B.14}$$

$$
L_2(t)=\begin{cases}
0\,, & -6\le t\le -3,\\[2ex]
t+3+\frac{1}{2\pi}\sin\left[\pi(2t+5)\right]+\frac{3}{\pi}\sin\left[\frac{\pi}{3}t\right] & \\
\quad +\frac{3}{10\pi}\left\{\sin\left[\pi(2t+4)\right]-\sin\left[\frac{\pi}{3}(t-3)\right]\right\} & \\
\quad +\frac{3}{14\pi}\left\{\sin\left[\pi(2t+6)\right]+\sin\left[\frac{\pi}{3}(t+3)\right]\right\}, & -3\le t\le -2,\\[2ex]
1+\frac{3}{\pi}\left\{\sin\left[\frac{\pi}{3}(1-t)\right]+\sin\left[\frac{\pi}{3}t\right]\right\} & \\
\quad +\frac{3}{10\pi}\left\{\sin\left[\frac{\pi}{3}(2+t)\right]+\sin\left[\frac{\pi}{3}(3-t)\right]\right\} & \\
\quad +\frac{3}{14\pi}\left\{\sin\left[\frac{\pi}{3}(4-t)\right]+\sin\left[\frac{\pi}{3}(3+t)\right]\right\}, & -2\le t\le 3,\\[2ex]
4-t+\frac{1}{2\pi}\sin\left[\pi(7-2t)\right]+\frac{3}{\pi}\sin\left[\frac{\pi}{3}(1-t)\right] & \\
\quad +\frac{3}{10\pi}\left\{\sin\left[\frac{\pi}{3}(t+2)\right]-\sin\left[\pi(2t-6)\right]\right\} & \\
\quad +\frac{3}{14\pi}\left\{\sin\left[\frac{\pi}{3}(4-t)\right]-\sin\left[\pi(2t-8)\right]\right\}, & 3\le t\le 4,\\[2ex]
0\,, & 4\le t\le 6,
\end{cases}
\tag{B.15}
$$

and

$$
L_3(t)=\begin{cases}
0\,, & -6\le t\le -2,\\[2ex]
t+2+\frac{1}{2\pi}\sin\left[\pi(2t+3)\right]+\frac{3}{\pi}\sin\left[\frac{\pi}{3}(t-1)\right] & \\
\quad +\frac{3}{10\pi}\left\{\sin\left[\pi(2t+2)\right]-\sin\left[\frac{\pi}{3}(t-4)\right]\right\} & \\
\quad +\frac{3}{14\pi}\left\{\sin\left[\pi(2t+4)\right]+\sin\left[\frac{\pi}{3}(t+2)\right]\right\}, & -2\le t\le -1,\\[2ex]
1+\frac{3}{\pi}\left\{\sin\left[\frac{\pi}{3}(2-t)\right]-\sin\left[\frac{\pi}{3}(1-t)\right]\right\} & \\
\quad +\frac{3}{10\pi}\left\{\sin\left[\frac{\pi}{3}(1+t)\right]-\sin\left[\frac{\pi}{3}(t-4)\right]\right\} & \\
\quad +\frac{3}{14\pi}\left\{\sin\left[\frac{\pi}{3}(5-t)\right]+\sin\left[\frac{\pi}{3}(2+t)\right]\right\}, & -1\le t\le 4,\\[2ex]
5-t-\frac{1}{2\pi}\sin\left[\pi(2t-9)\right]+\frac{3}{\pi}\sin\left[\frac{\pi}{3}(2-t)\right] & \\
\quad +\frac{3}{10\pi}\left\{\sin\left[\frac{\pi}{3}(t+1)\right]-\sin\left[\pi(2t-8)\right]\right\} & \\
\quad +\frac{3}{14\pi}\left\{-\sin\left[\frac{\pi}{3}(t-5)\right]-\sin\left[\pi(2t-10)\right]\right\}, & 4\le t\le 5,\\[2ex]
0\,, & 5\le t\le 6,
\end{cases}
\tag{B.16}
$$

The function $y(t)$ is plotted as the solid curve on the right in Figure 1. All of the details of the 3 peaks are so smeared together by the convolution with the kernel function that that there is no hint of any structure in the underlying $x(\xi)$.

C Discretizing the Modified Phillips Problem

The modified Phillips problem described in Appendix B was discretized by choosing $m = 300$ equally spaced mesh points on the interval $-5.9625 \leq t \leq 5.9625$ to give

$$y(t_i) = \int_{-3}^{3} K(t_i, \xi) x(\xi)\, d\xi \;, \quad i = 1, 2, \ldots, 300 \;, \tag{C.1}$$

and by replacing each of the integrals by an $n = 241$ point trapezoidal quadrature sum, i.e.,

$$\int_{-3}^{3} K(t_i, \xi) x(\xi) d\xi \approx \sum_{j=1}^{241} \omega_j K(t_i, \xi_j) x(\xi_j), \; i = 1, \ldots, 300 \;, \tag{C.2}$$

where

$$(\, \omega_1 \,,\, \omega_2 \,,\, \omega_3 \,,\, \ldots \,,\, \omega_{240} \,,\, \omega_{241} \,)^T = \frac{0.025}{2} \times (\, 1 \,,\, 2 \,,\, 2 \,,\, \ldots \,,\, 2 \,,\, 1 \,)^T \,. \tag{C.3}$$

Defining the n-vector $\mathbf{x}^*$ and the m-vector $\mathbf{y}^*$ by

$$\begin{array}{lcl} x_j^* = x(\xi_j) & , & j = 1 \,,\, 2 \,,\, \ldots \,,\, 241 \,, \\ y_i^* = y(t_i) & , & i = 1 \,,\, 2 \,,\, \ldots \,,\, 300 \,, \end{array} \tag{C.4}$$

and the $m \times n$ matrix $\mathbf{K}$ by

$$\begin{array}{lcl} K_{i,j} = \omega_j K(t_i, \xi_j) & , & i = 1 \,,\, 2 \,,\, \ldots \,,\, 300 \,, \\ & , & j = 1 \,,\, 2 \,,\, \ldots \,,\, 241 \,, \end{array} \tag{C.5}$$

gives

$$\mathbf{y}^* = \mathbf{K}\mathbf{x}^* + \boldsymbol{\delta} \,, \tag{C.6}$$

where $\boldsymbol{\delta}$ is an m-vector of quadrature errors. A crucial assumption in replacing the intregrals with quadrature sums is that the value of n is chosen large enough so that the δ_i are small relative to the random measuring errors ϵ_i. To assure that this assumption was satisfied for the test problem, the elements of the vector $\mathbf{y}^*$ were computed from the matrix-vector product

$$\mathbf{y}^* = \mathbf{K}\mathbf{x}^* \tag{C.7}$$

rather than from Eqs. (B.9) - (B.16). More precisely, the matrix elements $K_{i,j}$ were computed from Eqs. (B.2), (C.3), (C.5), the vector elements x_j^* were computed from Eqs. (B.3) - (B.6), and the vector $\mathbf{y}^*$ was then computed from Eq. (C.7). The "measured" vector $\mathbf{y}$ was then obtained by adding random perturbations to the elements of this $\mathbf{y}^*$. Each perturbation was chosen independently from a normal distribution with mean zero and standard deviation $s_i = (10^{-5})\sqrt{y_i^*}$, so the variance matrix was

$$\mathbf{S}^2 = \mathbf{diag}(s_1^2,\ s_2^2,\ \ldots,\ s_{300}^2)\ , \quad s_i = (10^{-5})\sqrt{y_i^*}\ . \tag{C.8}$$

D The Cumulative Periodogram

For a given time series, a periodogram is an estimate of the power spectrum of the series, i.e., an estimate of how the total variance in the series is distributed in frequency. Such an estimate can be obtained at any desired number of frequencies $f_k,\ k = 1, 2, \cdots, N$, in the interval $0 \leq f \leq \frac{1}{2T}$, where T is the sample spacing for the time variable, and N is chosen to be greater than or equal to the number of sample points in the time series. If the time series is taken to be the residuals r_i for an estimated solution to an ill-posed problem, with the element number i as the time variable, then $T = 1$, and the periodogram is estimated on an N-point equally spaced mesh on the interval $0 \leq f \leq 0.5$. It is obtained by zero-padding the r_i series to have $N \geq m$ terms and computing the discrete Fourier transform

$$\mathcal{R}_k = T \sum_{j=1}^{N} r_j \exp\left(-\,\mathrm{i}\, 2\pi \frac{kj}{N}\right)\ , \quad k = 0, 1, 2, \ldots, \frac{N}{2}\ , \tag{D.1}$$

where $\mathrm{i} = \sqrt{-1}$ and

$$r_j = \begin{cases} b_j - (\mathbf{A}\hat{\mathbf{x}})_j & ,\quad j = 1, 2, \ldots, m\ , \\ 0 & ,\quad j = m+1, m+2, \ldots, N\ . \end{cases} \tag{D.2}$$

The zero-padding increases the density of the frequency mesh, but does not change the value of the transform at any given frequency. Each $\mathcal{R}_k$ is associated with the corresponding Fourier frequency $f_k = \frac{k}{NT}$. The periodogram is computed from the transform by

$$\mathcal{P}_k = \mathcal{P}(f_k) = \frac{1}{NT}|\mathcal{R}_k|^2\ , \quad k = 0, 1, 2, \ldots, \frac{N}{2}\ . \tag{D.3}$$

Peaks in the plot of $\mathcal{P}_k$ versus f_k indicate the presence of sinusoidal cycles in the parent time series, and the height of each such peak gives an estimate of the power of that cycle.

The cumulative peridogram is defined by

$$\mathcal{C}_0 = \mathcal{C}(0) = 0 \; , \qquad \mathcal{C}_k = \mathcal{C}(f_k) = \frac{1}{S}\sum_{j=1}^{k} \mathcal{P}_j \; , \;\; k = 1, 2, \ldots, \frac{N}{2} \; , \tag{D.4}$$

where

$$S = \sum_{j=1}^{N/2} \mathcal{P}_j \; . \tag{D.5}$$

Clearly $\mathcal{C}(f)$ is a monotonic nondecreasing function of frequency ranging between the values $\mathcal{C}(0) = 0$ and $\mathcal{C}(1/2T) = 1$. The relationship between the periodogram and the cumulative periodogram is illustrated in Figures 12 and 13. In the first case the time series is comprised of $m = 128$ independent random samples x_i from a standard normal distribution (mean = 0, variance = 1), and in the second case the series is given by

$$x_i = \sin\left(\frac{2\pi i}{10}\right) + n_i \; , \;\; i = 1, 2, \ldots, 128 \; , \tag{D.6}$$

where the n_i are independent samples from a zero-mean normal distribution with variance = 0.1. Thus the first series is a realization of a pure white noise signal and the second is a single sinusoid, with unit amplitude and period 10, corrupted by zero-mean white noise with standard deviation 0.316. In both cases the time series was first detrended by subtracting out the mean value which was approximately, but not exactly, zero. The periodogram for the white noise case has the power distributed more or less uniformly on the interval $[0, 0.5]$, so the cumulative periodgram does not depart too much from a diagonal line that would connect the points $(0, 0)$ and $(0.5, 1)$, i.e., the line

$$\mathcal{C}(f) = 2Tf \; , \tag{D.7}$$

which is the theoretical distribution for pure white noise. One way to quantize departures from white noise is to compare the length of that line,

$$\sqrt{(0.5)^2 + (1.0)^2} = 1.11803 \; ,$$

with that of the computed estimate,

$$\text{Length}(\mathcal{C}) = \sum_{k=1}^{N/2} \left[\sqrt{(\mathcal{C}_k - \mathcal{C}_{k-1})^2 + (f_k - f_{k-1})^2} \; \right] \; , \tag{D.8}$$

which is given in the legend box on the plot of the cumulative periodogram.

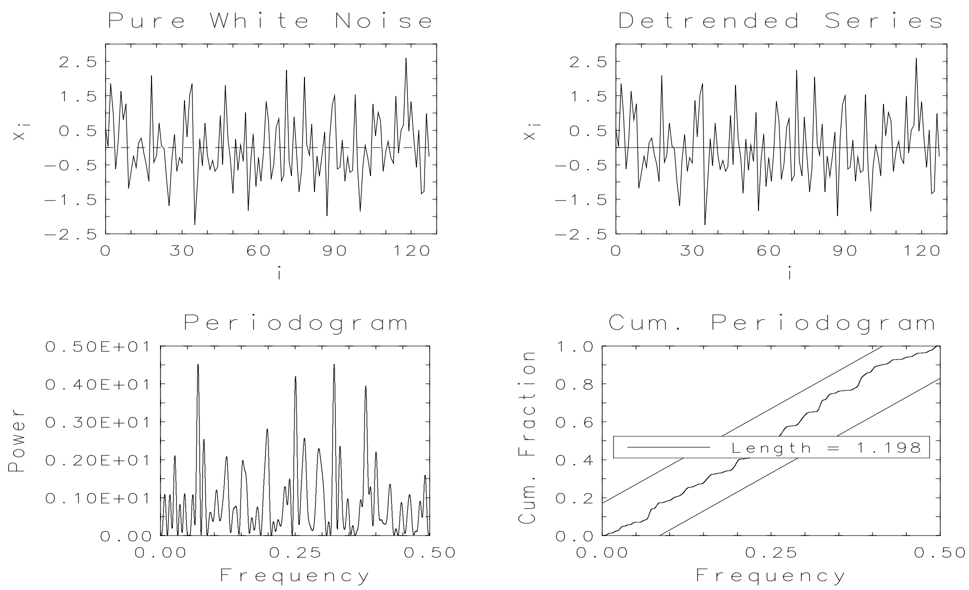

Figure 12: Periodogram and cumulative periodogram for a white noise time series

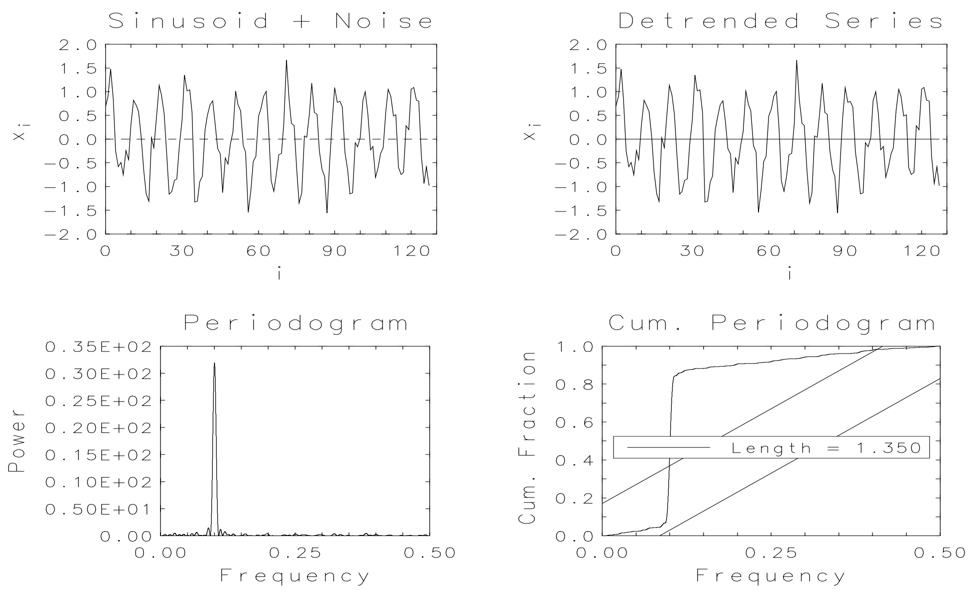

Figure 13: Periodogram and cumulative periodogram for a single sinusoid plus white noise time series

The periodogram for the sinusoid plus noise time series is dominated by a single peak at frequency 0.1 (period = 10) which produces a wide deviation from the white noise diagonal at the same frequency in the cumulative periodogram. One way to test whether of not such a deviation represents a statistically significant departure from white noise is to construct the two parallel off-diagonal lines defined by

$$\mathcal{C}(f) = \pm\delta + 2Tf \ , \tag{D.9}$$

where δ is the 5% point of the Kolmogorov-Smirnov statistic for a sample of size $m/2$. These two lines, which are plotted in both Figures 12 and 13, enclose a 95% confidence band which can be used to test the hypothesis that the time series is a realization of white noise. The cumulative periodogram ordinates for such a series should lie outside this band for at most 5% of the frequencies. None of the ordinates lie outside the band for the white noise case, but more than 50% of them do for the sinusoid plus noise case.

References

[1] Anderson, E., Bai, Z., Bischof, C., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., Ostrouchov, S. and Sorensen, D. (1992), *LAPACK Users' Guide*, SIAM, Philadelphia.

[2] Anderson, T.W. (1958), *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New York, Chapt. 2.

[3] Beyer, W.H. (1968), *Handbook of Tables for Probability and Statistics*, CRC Press, Boca Raton, pp. 293-294.

[4] Boisvert, R.F., Howe, S.E. and Kahaner, D.K. (1984), *The guide to available mathematical software (GAMS)*, PB 84-171305, National Technical Information Service, Springfield, VA.

[5] Burrus, W. and Verbinski, V. (1969), *Fast neutron spectroscopy with thick organic scintillators*, Nuclear Instruments and Methods, 67, pp.181-196.

[6] Fuller, W.A. (1976), *Introduction to Statistical Time Series*, John Wiley & Sons, New York, Chapt. 7.

[7] GAMS (1998), *Guide to Available Mathematical Software*, http://gams.nist.gov/.

[8] Golub, G. and Kahan, W. (1965), *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. Ser. B, 2, pp. 205-224.

[9] Golub, G., Klema, V. and Stewart, G.W. (1976), *Rank Degeneracy and Least Squares Problems*, Technical Report TR-456, Department of Computer Science, University of Maryland.

[10] Graybill, F.A. (1976), *Theory and Application of the Linear Model*, Duxbury Press, North Scituate, Chapt. 6.

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

[12] Hansen, P.C. (1998), *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, Chapt. 4.

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

[14] Hogg, R.V. and Craig, A.T. (1965), *Introduction to Mathematical Statistics*, Macmillan, New York, Sect. 4.6.

[15] Kahaner, D., Moler, C., and Nash, S. (1989), *Numerical Methods and Software*, Prentice Hall, Englewood Cliffs, Chapt. 3.

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

[17] Smithies, F. (1958), *Integral Equations*, Cambridge University Press, Cambridge, Chapt. 8.

[18] Tarantola, A. (1987), *Inverse Problem Theory*, Elsevier, Amsterdam, Chapt. 1.

[19] Twomey, S. (1977), *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*, Elsevier, Amsterdam, Chapt. 6.

[20] Verbinski, V., Burrus, W., Love, T., Zobel, W., Hill, N., and Textor, R. (1968), *Calibration of an organic scintillator for neutron spectrometry*, Nuclear Instruments and Methods, 65, pp. 8-25.

[21] Wing, G.M. (1991), *A Primer on Integral Equations of the First Kind*, SIAM, Philadelphia, Chapt. 2.