1-1-2009

# A Newton Root-Finding Algorithm For Estimating the Regularization Parameter For Solving Ill-Conditioned Least Squares Problems

Jodi Mead
*Boise State University*

Rosemary Renaut
*Arizona State University at the Tempe Campus*

# A Newton root-finding algorithm for estimating the regularization parameter for solving ill-conditioned least squares problems

**Jodi L. Mead**[1] **and Rosemary A. Renaut**[2]

[1]Supported by NSF grant EPS 0447689, Boise State University, Department of Mathematics, Boise, ID 83725-1555. Tel: 208426-2432, Fax: 208-426-1354. Email: jmead@boisestate.edu
[2]Supported by NSF grants DMS 0513214 and DMS 0652833. Arizona State University, School of Mathematical and Statistical Sciences, Tempe, AZ 85287-1804. Tel: 480-965-3795, Fax: 480-965-4160. Email: renaut@asu.edu

**Abstract.** We discuss the solution of numerically ill-posed overdetermined systems of equations using Tikhonov *a-priori*-based regularization. When the noise distribution on the measured data is available to appropriately weight the fidelity term, and the regularization is assumed to be weighted by inverse covariance information on the model parameters, the underlying cost functional becomes a random variable that follows a $\chi^2$ distribution. The regularization parameter can then be found so that the optimal cost functional has this property. Under this premise a scalar Newton root-finding algorithm for obtaining the regularization parameter is presented. The algorithm, which uses the singular value decomposition of the system matrix is found to be very efficient for parameter estimation, requiring on average about 10 Newton steps. Additionally, the theory and algorithm apply for Generalized Tikhonov regularization using the generalized singular value decomposition. The performance of the Newton algorithm is contrasted with standard techniques, including the L-curve, generalized cross validation and unbiased predictive risk estimation. This $\chi^2$-curve Newton method of parameter estimation is seen to be robust and cost effective in comparison to other methods, when white or colored noise information on the measured data is incorporated.

*Tikhonov regularization, least squares, regularization parameter*

AMS classification scheme numbers: 15A09, 15A29, 65F22, 62F15, 62G08

Submitted to: *Inverse Problems, 24 October 2008*

## 1. Introduction

We discuss the solution of numerically ill-posed, potentially overdetermined, systems of equations $A\mathbf{x} = \mathbf{b}$, $A \in \mathcal{R}^{m\times n}$, $\mathbf{b} \in \mathcal{R}^m$, $\mathbf{x} \in \mathcal{R}^n$, $m \geq n$. Such problems arise in many practical areas including image deblurring, for which the matrix $A$ represents the deconvolution operator, or the solution of Volterra integral equations. Extensive background information on these problems and their solution is available in standard references such as Hansen [8], Vogel [23] and the recent text of Hansen, Nagy, and O'Leary [9]. Because of the ill-conditioning of the matrix $A$, solutions cannot be found by straightforward solution of the least squares data-fit problem

$$\hat{\mathbf{x}}_{\rm ls} = \operatorname{argmin}\{\|A\mathbf{x} - \mathbf{b}\|_2\}. \tag{1}$$

Instead, relevant methods include projection-based techniques, which seek to project the noise out of the system and lead to solutions of reduced problems, [17], and algorithms in which a regularization term is introduced. Both directions introduce complications, not least of which are stopping criteria for the projection iterations in the first case and in the second case finding a suitable regularization parameter which trades-off a regularization term relative to the data-fitting or fidelity term (1). Here we focus on efficient and robust determination of the regularization parameter.

The most general formulation considered here is the regularized weighted least squares problem with a two norm regularization term:

$$\hat{\mathbf{x}}_{\rm rls} = \operatorname{argmin} J(\mathbf{x}) = \operatorname{argmin}\{\|A\mathbf{x} - \mathbf{b}\|^2_{W_\mathbf{b}} + \|D(\mathbf{x} - \mathbf{x}_0)\|^2_{W_\mathbf{x}}\}, \tag{2}$$

where the weighted norm is $\|\mathbf{y}\|^2_W = \mathbf{y}^T W\mathbf{y}$, for general vector $\mathbf{y}$ and weighting matrix $W$. Vector $\mathbf{x}_0$ is a given reference vector of *a priori* information for the unknown model parameters $\mathbf{x}$ and the matrix $D$ is typically chosen to yield approximations to the $l^{th}$ order derivative, $l = 0, 1, 2$, e.g. Chapter 1, [8]. In this case, except for $l = 0$ when $D = I_n$, the matrix $D \in \mathcal{R}^{(n-l)\times n}$ is necessarily not of full rank, but the invertibility condition

$$\mathcal{N}(A) \cap \mathcal{N}(D) \neq 0, \tag{3}$$

is assumed, where here $\mathcal{N}(A)$ is the null space of matrix $A$. If information on the noise structure of the measurements $\mathbf{b}$ is available, then $W_\mathbf{b}$ is taken to be the inverse of $C_\mathbf{b}$, the error covariance matrix for $\mathbf{b}$. For example, for colored noise the covariance is given by $C_\mathbf{b} = \operatorname{diag}(\sigma^2_{b_i})$, where $\sigma^2_{b_i}$ is the variance in the $i^{\rm th}$ component of $\mathbf{b}$, or for white noise $C_\mathbf{b} = \sigma^2 I_m$, where $\sigma^2$ is the common variance in the components of $\mathbf{b}$. Matrix $W_\mathbf{x}$ is generally

replaced by $\lambda^2 I_n$, where $\lambda$ is an unknown **regularization parameter**; its determination using statistical properties of the functional $J(\hat{\mathbf{x}}_{\rm rls})$ is the focus of this paper.

The outline of the paper is as follows. In Section 2 we present the theoretical background and development of our approach for obtaining the regularization parameter, based on existing results of [18] and [1]. The main new theorem, which characterizes the cost functional of (2) for arbitrary $D$, is presented in Section 3 and leads to the design of a Newton-based algorithm for estimating the regularization parameter. This algorithm, which as presented here, uses the gneralized singular value decomposition (GSVD) [3, 5], is also contrasted with other standard methods, including the L-curve, generalized cross validation (GCV) and unbiased predictive risk estimator (UPRE), [7, 23] in Section 4. Numerical results for both simulated data and a real data example from hydrology are presented in Section 5. Although our main focus here is on finding $W_{\mathbf{x}}$, given a reliable estimate of $W_{\mathbf{b}}$, we also briefly discuss the reverse situation, finding $W_{\mathbf{b}}$ given $W_{\mathbf{x}}$, in Sections 2.2 and 3.1. Indeed, the real data example from hydrology is posed in this alternative framework. Algorithmic details are reserved for the Appendix. Future work and conclusions are discussed in Section 6.

## 2. Theoretical Background

It is well-established, stated as *Rao's First Fundamental Theorem of Least Squares, 1973* (page 189 [18]), that, assuming that the matrix $A$ has rank $r$, the cost functional $J(\hat{\mathbf{x}}_{\rm ls})$ for the unregularized case (1), with weighting $W_{\mathbf{b}} = \sigma_{\mathbf{b}}^{-2} I_m$, is a random variable which follows a $\chi^2$ distribution with $m - r$ degrees of freedom, when the white noise $\epsilon_i$ in $b_i$ has common covariance $\sigma_{\mathbf{b}}^2 I_m$. Apparently less well-known is the extension of this result for the regularized problem (2) with matrix $D = I_n$. Specifically, when $W_{\mathbf{b}}$ and $W_{\mathbf{x}}$ are inverse covariance matrices on the mean zero errors in data $\mathbf{b}$ and initial parameter estimates $\mathbf{x}_0$, respectively, the functional $J(\hat{\mathbf{x}}_{\rm rls})$ in (2) is a random variable which follows a $\chi^2$ distribution with $m$ degrees of freedom [1, 15]. A further extension for general $D$ and a detailed proof are given in Section 3. These results extend the observation that $\hat{\mathbf{x}}_{\rm rls}$ is the maximum-likelihood estimate, e.g. [23, 8].

Mead [15] suggested capitalizing on the $\chi^2$ distribution of the optimal functional to find an appropriate regularization matrix $W_{\mathbf{x}}$. In particular, when statistical information on $\epsilon$ is available $W_{\mathbf{x}}$ may be chosen in order that $J(\hat{\mathbf{x}}_{\rm rls}(\mathrm{W}_{\mathbf{x}}))$ as closely as possible follows a $\chi^2$ distribution with mean $m$:

$$m - \sqrt{2m} z_{\alpha/2} < J(\hat{\mathbf{x}}_{\rm rls}(\mathrm{W}_{\mathbf{x}})) < m + \sqrt{2m} z_{\alpha/2}. \tag{4}$$

Here it is assumed that $m$ is large, in which case the $\chi^2$ distribution with $m$ degrees of freedom can be approximated by the normal distribution with mean $m$ and variance $2m$. Thus $z_{\alpha/2}$ is the relevant $z$-value for for a standard normal distribution, and $\alpha$ defines the $(1 - \alpha)$ confidence interval that $J$ is a $\chi^2$ random variable with $m$ degrees of freedom. For example,

when $m = 100$ the $95\%$ confidence interval states that $J(\hat{\mathbf{x}}_{\rm rls}(\mathrm{w}_{\mathbf{x}}))$ must lie in the interval $[72.28, 127.72]$.

To utilize this result with $D = I$, after a little algebra and using $C_{\mathbf{x}} = W_{\mathbf{x}}^{-1}$, it was shown in [15]

$$\hat{\mathbf{x}}_{\rm rls}(\mathrm{w}_{\mathbf{x}}) = \mathbf{x}_0 + C_{\mathbf{x}}A^T(AC_{\mathbf{x}}A^T + C_{\mathbf{b}})^{-1}\mathbf{r}, \quad \text{and} \tag{5}$$

$$J(\hat{\mathbf{x}}_{\rm rls}(\mathrm{w}_{\mathbf{x}})) = \mathbf{r}^T(AC_{\mathbf{x}}A^T + C_{\mathbf{b}})^{-1}\mathbf{r}, \tag{6}$$

where $\mathbf{r} = \mathbf{b} - A\mathbf{x}_0$. Therefore, assuming $C_{\mathbf{x}}$ is symmetric positive definite (SPD), so that its Cholesky factorization $L_{\mathbf{x}}L_{\mathbf{x}}^T$† exists, where $L_{\mathbf{x}}$ is lower triangular and invertible, Mead solved the problem of finding $L_{\mathbf{x}}$, and hence $C_{\mathbf{x}}$, through a nonlinear minimization.

**Algorithm 1 ([15])** *Given confidence interval parameter $\alpha$, initial residual $\mathbf{r} = \mathbf{b} - A\mathbf{x}_0$ and estimate of the data covariance $C_{\mathbf{b}}$, find $L_{\mathbf{x}}$ which solves the nonlinear optimization.*

$$\begin{array}{ll} \text{Minimize} & \|L_{\mathbf{x}}L_{\mathbf{x}}^T\|_F^2 \\ \text{Subject to} & m - \sqrt{2m}z_{\alpha/2} < \mathbf{r}^T(AL_{\mathbf{x}}L_{\mathbf{x}}^TA^T + C_{\mathbf{b}})^{-1}\mathbf{r} < m + \sqrt{2m}z_{\alpha/2} \\ & AL_{\mathbf{x}}L_{\mathbf{x}}^TA^T + C_{\mathbf{b}} \text{ well-conditioned.} \end{array}$$

Moreover, given estimates of $W_{\mathbf{b}}$ and $W_{\mathbf{x}}$ satisfying (4) for small $\alpha$, the posterior probability density for $\mathbf{x}$, $\mathcal{G}(\mathbf{x})$, is a Gaussian probability density such that

$$\mathcal{G}(\mathbf{x}) = B\ \exp(-\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}}_{\rm rls})^T\tilde{W}_{\mathbf{x}}(\mathbf{x} - \hat{\mathbf{x}}_{\rm rls})), \tag{7}$$

where $\tilde{W}_{\mathbf{x}}$ is the inverse posterior covariance probability density, and $B$ is a constant. In this case, it can be shown that

$$\tilde{W}_{\mathbf{x}} = A^TW_{\mathbf{b}}A + W_{\mathbf{x}}, \quad \tilde{C}_{\mathbf{x}} = (A^TW_{\mathbf{b}}A + W_{\mathbf{x}})^{-1} \tag{8}$$

[21]. This identification of the posterior covariance is then useful in assigning uncertainty bars to the posterior values of $\hat{\mathbf{x}}_{\rm rls}$, particularly when $\tilde{W}_{\mathbf{x}}$ is diagonal, [15]. While Mead [15] illustrated the use of Algorithm 1, and the associated posterior information (8), practically some refinements are needed to make the approach computationally feasible. Here we focus on the single variable case $W_{\mathbf{x}} = \lambda^2 I_n = \sigma_{\mathbf{x}}^{-2} I_n$.

### *2.1. Single Variable Case*

The singular value decomposition (SVD) facilitates the derivation of the algorithm for a single regularization parameter, $\lambda$. Assuming that $C_{\mathbf{b}}$ is SPD then its Choleski factorization $C_{\mathbf{b}} = L_{\mathbf{b}}L_{\mathbf{b}}^T$ exists, where $L_{\mathbf{b}}$ is invertible, and we use the SVD of $L_{\mathbf{b}}^{-1}A$.

† Note that throughout the paper we deliberately make use of the Cholesky factorization in all algebraic manipulations and simplifications with the covariance matrices, in keeping with practical numerical implementations.

**Lemma 2.1** *Let $U\Sigma V^T$ be the singular value decomposition (SVD) of $L_{\mathbf{b}}^{-1}A$, [5], with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$, where $r \leq \min\{m,n\}$ is the rank of matrix $A$. Then with $C_{\mathbf{x}} = L_{\mathbf{x}}L_{\mathbf{x}}^T$ (6) is replaced by*

$$m - \sqrt{2m}z_{\alpha/2} \quad < \mathbf{s}^T P^{-1}\mathbf{s} < m + \sqrt{2m}z_{\alpha/2}, \quad \mathbf{s} = U^T L_{\mathbf{b}}^{-1}\mathbf{r}, \tag{9}$$

$$P = \Sigma V^T L_{\mathbf{x}} L_{\mathbf{x}}^T V \Sigma^T + I_m, \tag{10}$$

*and the posterior covariance matrix is given by*

$$\tilde{C}_{\mathbf{x}} = L_{\mathbf{x}}(L_{\mathbf{x}}^T V \Sigma^T \Sigma V^T L_{\mathbf{x}} + I_n)^{-1} L_{\mathbf{x}}^T. \tag{11}$$

**Proof.** The following algebraic manipulations are standard, see for example [23].

$$\begin{aligned} AL_{\mathbf{x}}L_{\mathbf{x}}^T A^T + L_{\mathbf{b}}L_{\mathbf{b}}^T &= L_{\mathbf{b}}(L_{\mathbf{b}}^{-1}AL_{\mathbf{x}}L_{\mathbf{x}}^T A^T (L_{\mathbf{b}}^T)^{-1} + I_m)L_{\mathbf{b}}^T \\ &= L_{\mathbf{b}}(U\Sigma V^T L_{\mathbf{x}} L_{\mathbf{x}}^T V \Sigma^T U^T + UU^T)L_{\mathbf{b}}^T \quad \text{and} \\ \mathbf{r}^T(AL_{\mathbf{x}}L_{\mathbf{x}}^T A^T + L_{\mathbf{b}}L_{\mathbf{b}}^T)^{-1}\mathbf{r} &= \mathbf{r}^T(U^T L_{\mathbf{b}}^T)^{-1}(\Sigma V^T L_{\mathbf{x}} L_{\mathbf{x}}^T V \Sigma^T + I_m)^{-1}(L_{\mathbf{b}}U)^{-1}\mathbf{r} \\ &= \mathbf{s}^T(\Sigma V^T L_{\mathbf{x}} L_{\mathbf{x}}^T V \Sigma^T + I_m)^{-1}\mathbf{s}. \end{aligned}$$

The result for the covariance matrix follows similarly. ■

To be consistent with our use of $W_{\mathbf{x}}$ as an inverse covariance matrix, we now specifically set $C_{\mathbf{x}} = \sigma_{\mathbf{x}}^2 I_n$, equivalently throughout we now use $\sigma_{\mathbf{x}} = 1/\lambda$. Then the matrix $P$ in (10) simplifies and (9) becomes

$$m - \sqrt{2m}z_{\alpha/2} < \mathbf{s}^T \mathrm{diag}(\frac{1}{\sigma_i^2\sigma_{\mathbf{x}}^2 + 1})\mathbf{s} < m + \sqrt{2m}z_{\alpha/2}, \quad \sigma_i = 0, i > r. \tag{12}$$

This suggests that $\sigma_{\mathbf{x}}$ can be found by a single-variable root-finding Newton method to solve

$$F(\sigma_{\mathbf{x}}) = \mathbf{s}^T \mathrm{diag}(\frac{1}{1 + \sigma_{\mathbf{x}}^2\sigma_i^2})\mathbf{s} - m = 0, \tag{13}$$

within some tolerance related to the confidence interval parameter $\alpha$. Indeed, it is evident from (12) that $F(\sigma_{\mathbf{x}})$ is only required to lie in the interval $[-\sqrt{2m}z_{\alpha/2}, \sqrt{2m}z_{\alpha/2}]$, and this can be used to determine the tolerance for the Newton method. Specifically, setting $\tilde{m} = m - \sum_{i=r+1}^m s_i^2$ and introducing the vector

$$\tilde{s}_i = \frac{s_i}{\sigma_{\mathbf{x}}^2\sigma_i^2 + 1} \quad i = 1 \ldots r, \quad \text{and } 0 \quad i > r, \tag{14}$$

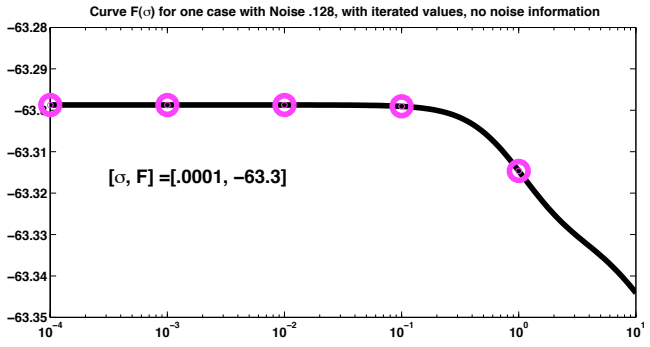
so as to reduce computation, we immediately obtain

$$F(\sigma_{\mathbf{x}}) = \mathbf{s}^T\tilde{\mathbf{s}} - \tilde{m} \quad \text{and} \tag{15}$$

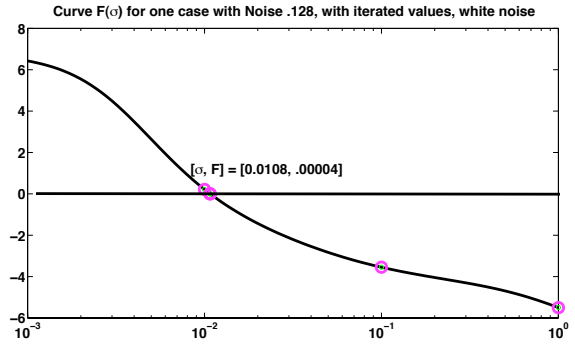
$$F'(\sigma_{\mathbf{x}}) = -2\sigma_{\mathbf{x}}\|\mathbf{t}\|_2^2, \quad t_i = \sigma_i\tilde{s}_i. \tag{16}$$

Illustrative examples for $F$, for one of the simulated data sets used in Section 5, are shown in Figure 1. We see that if $F$ has a positive root, then it is unique:

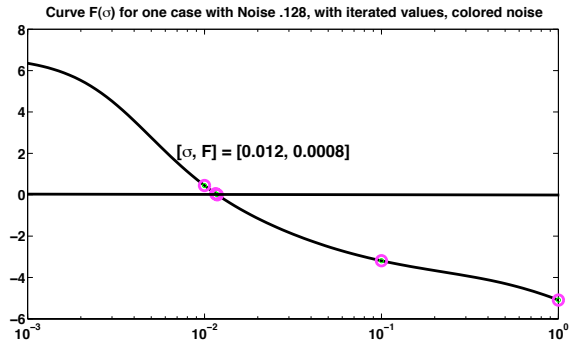
**Lemma 2.2** *Any positive root of (13) is unique.*




(a)




(b)




(c)

**Figure 1.** Illustration of $\chi^2$ curve for problem `phillips`, for one case with noise level .128, (a) no noise assumption made $C_{\mathbf{b}} = I_m$, (b) white noise and (c) colored noise. In each case the $x-$axis is the value of $\sigma_{\mathbf{x}}$ on a logarithmic scale and the $y-$axis is the value of the functional $F(\sigma_{\mathbf{x}})$ as given by (16).

**Proof.** From (16) and (13) it is immediate that $F(\sigma)$ is continuous, even and monotonically decreasing on $[0, \infty)$. There are just three possibilities; (i) $F$ is asymptotically greater than zero for all $\sigma > 0$, (ii) $F(0) < 0$ and no root exists, or there exists a unique positive root. ■

The cases for which $F(\sigma_{\mathbf{x}}) \neq 0$ are informative. Case (i), when $F(0) > 0$ and $\lim_{\sigma_{\mathbf{x}} \to \infty} F' > \sqrt{2m} z_{\alpha/2}$, effectively suggests that the number of degrees of freedom is exceeded for large $\sigma$. But this corresponds to small regularization parameter $\lambda$ meaning that the system does not require regularization and this case is thus not of practical interest. But in case (ii) with $F(0) = \|\mathbf{s}\|^2 - m < -\sqrt{2m} z_{\alpha/2}$, the monotonicity of $F$ assures that no solution of (13) exists, see Figure 1(a). Effectively, in this case the lack of a solution within a reasonable interval means that $J(\hat{\mathbf{x}}_{\text{rls}})$ does not follow the properties of a $\chi^2$ distribution with $m$ degrees of freedom; the number of degrees of freedom $m$ is too large for the given data, and weighting matrix $W_{\mathbf{b}}$ was not correctly approximated. For example, the case in Figure 1(a) uses an unweighted fidelity term, $W_{\mathbf{b}} = I_m$, corresponding to a situation where no statistical information is included and the statistical properties do not apply. Therefore, in determining whether a positive unique root exists, it suffices to test the value of $F$ for both $0$ and a large $\sigma$ to determine whether one of the first two situations occurs.

Equipped with these observations a Newton algorithm in which we first bracket the root can be implemented. Then, the standard Newton update given by

$$\sigma_{\mathbf{x}}^{\text{new}} = \sigma_{\mathbf{x}} + \frac{1}{2\sigma_{\mathbf{x}} \|\mathbf{t}\|_2^2} (\mathbf{s}^T \tilde{\mathbf{s}} - \tilde{m}). \tag{17}$$

will lead to this unique positive root. The algorithm is extended for the case of general operator $D$ replacing the identity in Section 3.1, and algorithmic details, including a line search, are presented in the Appendix.

At convergence the update and covariance matrix are immediately given by

$$\hat{\mathbf{x}}_{\text{rls}} = \mathbf{x}_0 + V\tilde{\mathbf{t}}, \quad \tilde{t} = \sigma_{\mathbf{x}}^2 \mathbf{t} \quad \text{and} \tag{18}$$

$$\tilde{C}_{\mathbf{x}} = \sigma_{\mathbf{x}}^2 (V \text{diag}(\frac{1}{\sigma_{\mathbf{x}}^2 \sigma_i^2 + 1}, I_{n-r}) V^T). \tag{19}$$

Moreover, as noted before, at convergence (or termination of the algorithm) $|F(\sigma_{\mathbf{x}})| = \tau$, for some tolerance $\tau = \sqrt{2m} z_{\alpha/2}$. Equivalently, $F^2 = 2m z_{\alpha/2}^2$ and the confidence interval $(1-\alpha)$ can be calculated. The larger the value of $\tau$, the larger the $(1 - \alpha)$ confidence interval and the greater the chance that any random choice of $\sigma_{\mathbf{x}}$ will allow $J(\hat{x})$ to fall in the appropriate interval. Equivalently, if $\tau$ is large, we have less confidence that $\sigma_{\mathbf{x}}$ is a good estimate of the standard deviation of the error in $\mathbf{x}_0$.

### *2.2. Estimating the data error*

The problem of parameter estimation can be posed with respect $C_{\mathbf{x}}$ known and $C_{\mathbf{b}}$ to be found. The analysis is equivalent and thus not given. Hence suppose $C_{\mathbf{x}}$, possibly dense, is given and

$C_{\mathbf{b}} = \sigma_{\mathbf{b}}^2 I_m$ is to be found. Then we have that

$$m - \sqrt{2m} z_{\alpha/2} < \mathbf{r}^T (U\Sigma\Sigma^T U^T + L_{\mathbf{b}} L_{\mathbf{b}}^T)^{-1} \mathbf{r} < m + \sqrt{2m} z_{\alpha/2}, \tag{20}$$

where now $U\Sigma V^T$ is the SVD of $AL_{\mathbf{x}}$. When $L_{\mathbf{b}} = \sigma_{\mathbf{b}} I_m$, let

$$F(\sigma_{\mathbf{b}}) = \mathbf{s}^T \mathrm{diag}(\frac{1}{\sigma_i^2 + \sigma_{\mathbf{b}}^2})\mathbf{s} - m, \quad \mathbf{s} = U^T \mathbf{r}, \tag{21}$$

where $\sigma_i$ are the singular values of the matrix $AL_{\mathbf{x}}$, $\sigma_i = 0, i > r$. Then $\sigma_{\mathbf{b}}$ can be found by applying a similar Newton's iteration to find the root of $F = 0$, see the Appendix.

## 3. Extension to Generalized Tikhonov Regularization

We now return to (2) for the case of general operator $D$ and obtain a result on the degrees of freedom in the functional through use of a decomposition which extends the SVD to a common factorization when one has a pair of matrices, the generalized singular value decomposition (GSVD), eg Chapter 2 [8], Chapter 8 [5] and research papers, [3, 6, 7].

**Lemma 3.1** *[5] Assume the invertibility condition (3) and $m \geq n \geq p = n - l$. There exist unitary matrices $U \in \mathcal{R}^{m\times m}$, $V \in \mathcal{R}^{p\times p}$, and a nonsingular matrix $X \in \mathcal{R}^{n\times n}$ such that*

$$A = U\tilde{\Upsilon}X^T, \quad D = V\tilde{M}X^T, \tag{22}$$

*where*

$$\tilde{\Upsilon} = \begin{bmatrix} \Upsilon & 0 \\ 0 & I_{n-p} \\ 0 & 0 \end{bmatrix}, \quad \Upsilon = \mathrm{diag}(\upsilon_1, \ldots, \upsilon_p) \in \mathcal{R}^{p\times p},$$
$$\tilde{M} = \begin{bmatrix} M, & 0_{p\times(n-p)} \end{bmatrix}, \quad M = \mathrm{diag}(\mu_1, \ldots, \mu_p) \in \mathcal{R}^{p\times p},$$

*and such that*

$$\begin{aligned} 0 \leq \upsilon_1 \leq \ldots \leq \upsilon_p \leq 1, \quad 1 \geq \mu_1 \geq \ldots \geq \mu_p > 0, \\ \upsilon_i^2 + \mu_i^2 = 1, \quad i = 1, \ldots p. \end{aligned} \tag{23}$$

**Theorem 3.1** *Suppose $C_{\mathbf{b}} = W_{\mathbf{b}}^{-1}$ is the SPD covariance matrix on the mean zero normally-distributed data error, $\epsilon_i$ and $C_{\mathbf{D}}$ is the rank deficient symmetric positive semi-definite covariance matrix for the model errors $\zeta_i = (\hat{\mathbf{x}}_{\mathrm{rls}} - \mathbf{x}_0)_i$, with $W_{\mathbf{D}} = D^T W_{\mathbf{x}} D$ its conditional inverse, where $W_{\mathbf{x}}$ is SPD, satisfying the Moore-Penrose conditions $W_{\mathbf{D}} C_{\mathbf{D}} W_{\mathbf{D}} = W_{\mathbf{D}}$ and $C_{\mathbf{D}} W_{\mathbf{D}} C_{\mathbf{D}} = C_{\mathbf{D}}$, and that the invertibility condition (3) holds. Then for large $m$ the minimum value of the functional $J$ is a random variable which follows a $\chi^2$ distribution with $m - n + p$ degrees of freedom.*

**Proof.** The solution of (2) with general operator $D$ is given by the solution of the normal equations

$$\begin{aligned} \hat{\mathbf{x}}_{\mathrm{rls}} &= \mathbf{x}_0 + (A^T W_{\mathbf{b}} A + D^T W_{\mathbf{x}} D)^{-1} A^T W_{\mathbf{b}} \mathbf{r} \qquad (24) \\ &= \mathbf{x}_0 + R(\mathrm{W}_{\mathbf{x}}) W_{\mathbf{b}}^{1/2} \mathbf{r}, \text{ where} \\ R(\mathrm{W}_{\mathbf{x}}) &= (A^T W_{\mathbf{b}} A + D^T W_{\mathbf{x}} D)^{-1} A^T W_{\mathbf{b}}^{1/2}, \quad \mathbf{r} = \mathbf{b} - A\mathbf{x}_0. \qquad (25) \end{aligned}$$

The minimum value of $J(\mathbf{x})$ in (2) is

$$\begin{aligned} J(\hat{\mathbf{x}}_{\text{rls}}) &= \mathbf{r}^T(ARW_{\mathbf{b}}^{1/2} - I_m)^T W_{\mathbf{b}}(ARW_{\mathbf{b}}^{1/2} - I_m)\mathbf{r} + \\ &\qquad \mathbf{r}^T(DRW_{\mathbf{b}}^{1/2})^T W_{\mathbf{x}}(DRW_{\mathbf{b}}^{1/2})\mathbf{r} \\ &= \mathbf{r}^T W_{\mathbf{b}}^{1/2}(I_m - W_{\mathbf{b}}^{1/2}AR(\,\mathrm{W}_{\mathbf{x}}))W_{\mathbf{b}}^{1/2}\mathbf{r} \\ &= \mathbf{r}^T W_{\mathbf{b}}^{1/2}(I_m - A(\,\mathrm{W}_{\mathbf{x}}))W_{\mathbf{b}}^{1/2}\mathbf{r}, \end{aligned} \tag{26}$$

where here $A(\,\mathrm{W}_{\mathbf{x}}) = W_{\mathbf{b}}^{1/2}AR(\,\mathrm{W}_{\mathbf{x}})$ is the influence matrix [23]. Using the GSVD for the matrix pair $[\tilde{A}, \tilde{D}]$, $\tilde{A} = W_{\mathbf{b}}^{1/2}A$ and $\tilde{D} = W_{\mathbf{x}}^{1/2}D$, to simplify matrix $A(\,\mathrm{W}_{\mathbf{x}})$ it can be seen that

$$I_m - A(\,\mathrm{W}_{\mathbf{x}}) = I_m - U\tilde{\Upsilon}\tilde{\Upsilon}^T U^T = U(I_m - \tilde{\Upsilon}\tilde{\Upsilon}^T)U^T.$$

Therefore

$$\begin{aligned} J(\hat{\mathbf{x}}_{\text{rls}}) &= \mathbf{r}^T W_{\mathbf{b}}^{1/2} U(I_m - \tilde{\Upsilon}\tilde{\Upsilon}^T)U^T W_{\mathbf{b}}^{1/2}\mathbf{r} \\ &= \sum_{i=1}^{p} \mu_i^2 s_i^2 + \sum_{i=n+1}^{m} s_i^2, \quad \mathbf{s} = U^T W_{\mathbf{b}}^{1/2}\mathbf{r}, \\ &= \sum_{i=1}^{p} k_i^2 + \sum_{i=n+1}^{m} k_i^2, \quad \mathbf{k} = QU^T W_{\mathbf{b}}^{1/2}\mathbf{r}, \end{aligned}$$

where

$$Q = \begin{bmatrix} M & 0 & 0 \\ 0 & I_{n-p} & 0 \\ 0 & 0 & I_{m-n} \end{bmatrix}.$$

It remains to determine whether the components are independently normally distributed variables with mean $0$ and variance $1$, i.e. whether $\mathbf{k}$ is a standard normal vector, which then implies that $J$ is a random variable following a $\chi^2$ distribution, [18].

By the assumptions on the data and model errors $\epsilon_i$ and $\zeta_i$ and standard results on the distribution of linear combinations of random variables, the components of $\mathbf{b} = A\mathbf{x}$ are normally distributed random variables and $\mathbf{b}$ has mean $A\mathbf{x}_0$ and covariance $C_{\mathbf{b}} + AC_{\mathbf{D}}A^T$. Therefore the residual $\mathbf{r} = \mathbf{b} - A\mathbf{x}_0$, where $\mathbf{x}_0$ is constant, is a random variable with mean $0$ and covariance $C_{\mathbf{b}} + AC_{\mathbf{D}}A^T$. Further, $W_{\mathbf{b}}^{1/2}\mathbf{r}$ has mean $0$ and covariance $I_m + \tilde{A}C_{\mathbf{D}}\tilde{A}^T$. Now, using the GSVD, we can write $C_{\mathbf{D}} = (X^T)^{-1}\text{diag}(M^{-2}, 0_{n-p})X^{-1}$, and thus

$$\begin{aligned} I_m + \tilde{A}C_{\mathbf{D}}\tilde{A}^T &= U(I_m + \tilde{\Upsilon}\text{diag}(M^{-2}, 0_{n-p})\tilde{\Upsilon}^T)U^T \\ &= U\begin{bmatrix} M^{-2} & 0 & 0 \\ 0 & I_{n-p} & 0 \\ 0 & 0 & I_{m-n} \end{bmatrix}U^T = UQ^{-2}U^T. \end{aligned}$$

Hence $\mathbf{k}$ has mean $0$ and variance $QU^T(I_m + \tilde{A}C_{\mathbf{D}}\tilde{A}^T)UQ = I_m$. Equivalently, $J$ is a sum of $m - n + p$ squared independent standard normal random variables $k_i$. The result follows. ■

When matrix $D$ is full rank and all other assumptions remain it is clear that this proof also provides a complete proof that $J(\hat{\mathbf{x}}_{\text{rls}})$ is a $\chi^2$ random variable with $m$ degrees of freedom.

Again the optimal weighting matrix $W_{\mathbf{x}}$ for the generalized Tikhonov regularization should satisfy, as in (4), with $\tilde{\mathbf{r}} = W_{\mathbf{b}}^{1/2}\mathbf{r}$,

$$m-n+p-\sqrt{2(m-n+p)}z_{\alpha/2} < \tilde{\mathbf{r}}^T(I_m - A(\mathtt{W}_{\mathbf{x}}))\tilde{\mathbf{r}} < m-n+p+\sqrt{2(m-n+p)}z_{\alpha/2}.$$

Assuming that $W_{\mathbf{x}}$ has been found such that confidence in the hypotheses of Theorem 3.1 is high, the posterior probability density for $\mathbf{x}$, is given by (7) with

$$\tilde{W}_{\mathbf{x}} = A^T W_{\mathbf{b}} A + D^T W_{\mathbf{x}} D, \quad \tilde{C}_{\mathbf{x}} = (A^T W_{\mathbf{b}} A + D^T W_{\mathbf{x}} D)^{-1}. \tag{27}$$

Moreover, the development of an appropriate root finding algorithm in the single variable case follows as before, see the Appendix.

### *3.1. Single Variable Case*

Using the GSVD now for the pair $[\tilde{A}, D]$, it is immediate that

$$\begin{aligned} J(\hat{\mathbf{x}}_{\rm rls}) &= \tilde{\mathbf{r}}^T(I_m - \tilde{A}\tilde{C}_{\mathbf{x}}\tilde{A}^T)\tilde{\mathbf{r}}, \quad \tilde{C}_{\mathbf{x}} = (A^T W_{\mathbf{b}} A + \sigma_{\mathbf{x}}^{-2} D^T D)^{-1} \\ &= \mathbf{s}^T \begin{bmatrix} I_p - \Upsilon^2(\Upsilon^2 + \sigma_{\mathbf{x}}^{-2} M^2)^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{m-n} \end{bmatrix} \mathbf{s}, \quad \mathbf{s} = U\tilde{\mathbf{r}} \\ &= \sum_{i=1}^{p}(1 - \frac{\upsilon_i^2}{\upsilon_i^2 + \sigma_{\mathbf{x}}^{-2}\mu_i^2})s_i^2 + \sum_{i=n+1}^{m} s_i^2, \\ &= \sum_{i=1}^{p}(\frac{1}{\sigma_{\mathbf{x}}^2\gamma_i^2 + 1})s_i^2 + \sum_{i=n+1}^{m} s_i^2, \quad \gamma_i = \frac{\upsilon_i}{\mu_i}. \end{aligned} \tag{28}$$

As in Section 2.1, some algebra can be applied to reduce the computation for the calculation of $F$ and $F'$. Define $\tilde{m} = m - n + p - \sum_{i=1}^{p} s_i^2 \delta_{\gamma_i 0} - \sum_{i=n+1}^{m} s_i^2$, where $\delta_{\gamma_i,0}$ is the Kronecker delta, which is zero unless $\gamma_i = 0$, and let $\tilde{\mathbf{s}}$ be the vector of length $m$ with zero entries except $\tilde{s}_i = s_i/(\gamma_i^2\sigma_{\mathbf{x}}^2 + 1)$, $i = 1, \ldots, p$, for $\gamma_i \neq 0$, and let $t_i = \tilde{s}_i\gamma_i$. Then

$$F(\sigma_{\mathbf{x}}) = \mathbf{s}^T\tilde{\mathbf{s}} - \tilde{m}, \quad F'(\sigma_{\mathbf{x}}) = -2\sigma_{\mathbf{x}}\|\mathbf{t}\|_2^2, \tag{29}$$

and Lemma 2.2 still applies. Moreover, the Newton update (17) and the algorithms are the same but with $\sigma_i$ replaced everywhere by $\gamma_i$ in the appropriate definitions of the variables, see the Appendix. Additionally,

$$\hat{\mathbf{x}}_{\rm rls} = \mathbf{x}_0 + (X^T)^{-1}\tilde{\mathbf{t}}, \quad \tilde{t}_i = \begin{cases} \sigma_{\mathbf{x}}^2 t_i/\mu_i & i = 1 \ldots p \\ s_i & i = p+1 \ldots n. \end{cases} \tag{30}$$

$$\tilde{C}_{\mathbf{x}} = ((X^T)^{-1}\mathrm{diag}(\frac{\sigma^2}{\sigma^2\nu_i^2 + \mu_i^2}, I_{n-p})X^{-1}). \tag{31}$$

As for the SVD case, the situation in which information on the weighting of the operator $W_{\mathbf{x}}$ is known and $C_{\mathbf{b}}$ is to be estimated can also be considered. Note, also, that when $D$ is nonsingular and square, the GSVD of the matrix pair $[A, D]$, corresponds to the SVD of matrix $AD^{-1}$, with the singular values now ordered in the opposite direction. If either $A$ or

$D$ is ill-conditioned the calculation of $AD^{-1}$ can be unstable, leading to contamination of the results, and hence the algorithm should be, in general, implemented using the GSVD.

## 4. Other Parameter Estimation Techniques: L-curve, GCV and UPRE

The Newton update (17) for finding the regularization parameter based on root finding for the $\chi^2$-curve provides an alternative approach to standard techniques such as the L-curve, GCV and UPRE, which are also based on the use of the GSVD, for estimating the single variable regularization parameter. These methods are well-described in a number of research monographs and therefore no derivation is provided here. For example, see Chapter 7 [8] for a discusion on the L-curve and GCV which are implemented in the toolbox [7], Chapter 7 in [23] where GCV, L-curve and UPRE are discussed and the more recent text of [9] in which GCV, L-curve and discrepancy principle are again presented with some implementation details provided. We thus quote results from these texts, using notation in keeping with that used for the $\chi^2$ method. Moreover, in the following formulae the weighted right hand side $\tilde{\mathbf{b}}$ is replaced by the relevant weighted residual $\tilde{\mathbf{r}}$ when $\mathbf{x}_0 \neq 0$. In terms of the GSVD and using the notation of Section 3.1, the GCV function to be minimized is given by

$$\begin{aligned} C(\sigma) &= \frac{\|\tilde{\mathbf{b}} - \tilde{A}\mathbf{x}(\sigma)\|_2^2}{[\mathrm{trace}(I_m - A(\,\mathrm{W}_{\mathbf{x}}))]^2}, \quad W_{\mathbf{x}} = \sigma^{-2} I_n \\ &= \frac{\sum_{i=n+1}^{m} s_i^2 + \sum_{i=1}^{p} (\delta_{\gamma_i 0} s_i^2 + \tilde{s}_i^2)}{(m - n + (\sum_{i=1}^{p} \frac{1}{\gamma_i^2 \sigma^2 + 1}))^2}, \end{aligned} \tag{32}$$

[7]. The UPRE seeks to minimize the expected value of the predictive risk by finding the minimum of the UPRE function

$$\begin{aligned} U(\sigma) &= \|\tilde{\mathbf{b}} - \tilde{A}\mathbf{x}(\sigma)\|_2^2 + 2\,\mathrm{trace}(A(\,\mathrm{W}_{\mathbf{x}})) - m \\ &= (\sum_{i=n+1}^{m} s_i^2 + \sum_{i=1}^{p} (\delta_{\gamma_i 0} s_i^2 + \tilde{s}_i^2)) + 2(n - \sum_{i=1}^{p} \frac{1}{\gamma_i^2 \sigma^2 + 1}) - m, \end{aligned} \tag{33}$$

where note that the variance of the model error is explicitly included within the weighted residual, [23]. In contrast, the L-curve approach seeks to find the corner point of the plot, on log-log scale, of $\|D\mathbf{x}(\sigma)\|$ against $\|\tilde{A}\mathbf{x}(\sigma) - \tilde{\mathbf{b}}\|$. The advantages and disadvantages of these approaches are well noted in the literature. The L-curve does not yield optimal results for the weighted case $C_{\mathbf{b}} \neq I$, while the GCV and UPRE functions may be nearly flat for the optimal choice of $\sigma$, and/or have multiple minima, which thus leads to difficulty in finding the optimal argument, [7, 9]. See for example, Figure 7.2, p. 104 [23] which shows a relatively flat GCV curve, similar to what we also observed with some of our own experiments using GCV.

Another well-known method, which assumes white noise in the data, is the discrepancy principle. It is implemented by a Newton method, [23], and finds the variance $\sigma_{\mathbf{x}}^2$ such that

the regularized residual satisfies

$$\sigma_{\mathbf{b}}^2 = \frac{1}{m}\|\mathbf{b} - A\mathbf{x}(\sigma)\|_2^2. \tag{34}$$

Consistent with our notation this becomes the requirement that

$$\sum_{i=1}^{p}(\frac{1}{\gamma_i^2\sigma^2+1})^2 s_i^2 + \sum_{i=n+1}^{m} s_i^2 = m, \tag{35}$$

similar to (28), but note that the weight in the first sum is squared in this case. Because this method tends to lead to solutions which are over smoothed, [23], we do not use this method in the comparisons presented in the following section.

In each of these cases, the algorithms rely on multiple calculations of the regularization parameter. In particular, even though the GCV and UPRE functionals (32, 33) could be directly minimized, the optimal value is typically found by first evaluating the functional for a range of parameter values, on logarithmic scale. Then, after isolating a potential region for a minimum, this minimum is found within that range of parameter values. For small-scale problems, as considered in this paper, the parameter search is made more efficient by employing the GSVD (resp. SVD) for evaluating the relevant functions (32, 33), or for the data required for the L-curve. Effectively, in each case the GSVD is used to find solutions for at least $100$ different choices of the parameter value, [7]. This contrasts with the presented Newton method described in Sections 2.1, 3.1, which, as will be demonstrated in Section 5, converges with very few function evaluations. Hence while the cost of the proposed algorithm, as presented here, is dominated by that for obtaining the GSVD (resp. SVD), this cost is the same as that for initializing the other standard methods for parameter estimation, while in contrast the optimal parameter is found with minimal additional cost.

## 5. Experiments

For validation of the algorithm against other standard approaches we present a series of representative results using benchmark cases, `phillips`, `shaw`, `ilaplace` and `heat` from the Regularization Tool Box [7]. In addition, we present the results for a real model from hydrology. These experiments contrast the results presented in [15], in which Algorithm 1 was used with more general choices for the weighting matrix $W_{\mathbf{x}}$.

### *5.1. Benchmark Problems: Experimental Design*

System matrices $A$, right hand side data $\mathbf{b}$ and solutions $\mathbf{x}$ are obtained for a specific benchmark problem from the Regularization Tool Box [7]. In all cases we generate a random matrix $\Theta$ of size $m \times 500$, with columns $\Theta^c$, $c = 1:500$, using the Matlab®[13] function **randn** which generates variables from a standard normal distribution. Then setting $\mathbf{b}^c = \mathbf{b} + \text{level}\|\mathbf{b}\|_2\Theta^c/\|\Theta^c\|_2$, for $c = 1:500$, generates $500$ copies of the right hand vector

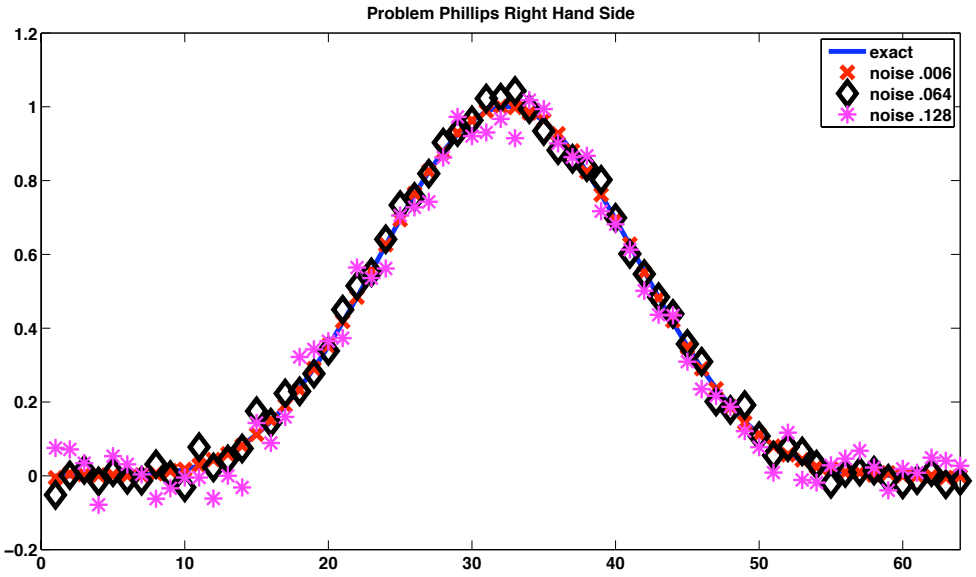



**Figure 2.** Illustration of the noise in the right hand side for problem `phillips`, for the three noise levels used in the presented experiments. The graphs show in each case $\mathbf{b}$ plotted against its index on the $x-$axis.

$\mathbf{b}$ with normally distributed noise, dependent on the chosen level. Results are presented for level $= .005, .05,$ and $.1$. An example of the error distribution for one case of **phillips** with $n = 64$ is illustrated in Figure 2. Because the noise depends on the right hand side $\mathbf{b}$ the actual errors, as measured by the mean of $\|\mathbf{b} - \mathbf{b}^c\|_\infty / \|\mathbf{b}\|_\infty$ over all $c$, are $.006$, $.064$ and $.128$, respectively.

To obtain the weighting matrix $W_\mathbf{b}$, the resulting covariance $C_\mathbf{b}$ between the measured components is calculated directly for the entire data set $\mathbf{B}$ with rows $(\mathbf{b}^c)^T$. Because of the design, $C_\mathbf{b}$ is close to diagonal, $C_\mathbf{b} \approx \mathrm{diag}(\sigma_{b_i}^2)$ and the noise is colored, see for example Figure 3, again for the same three noise distributions. For experiments assuming a white noise distribution the common variance $\sigma_\mathbf{b}^2$ is taken as the average of the $\sigma_{b_i}^2$. In all experiments, regardless of parameter selection method, the same covariance matrix is used.

The *a priori* reference solution $\mathbf{x}_0$ is generated using the exact known solution and noise added with level $= .1$ in the same way as for modifying $\mathbf{b}$. The same reference solution $\mathbf{x}_0$, see Figure 4, is used for all right hand side vectors $\mathbf{b}^c$, and for all algorithms, L-curve, GCV, UPRE and $\chi^2$.

### *5.2. Benchmark Problems: Results*

For all the tables the regularization for `shaw` uses the identity, while for the other problems the first derivative operator is used. The column $cb$ indicates white or colored noise assumption for matrix $C_\mathbf{b}$, $cb = 2, 3$, resp. In the implementation of the Newton method to find the optimal

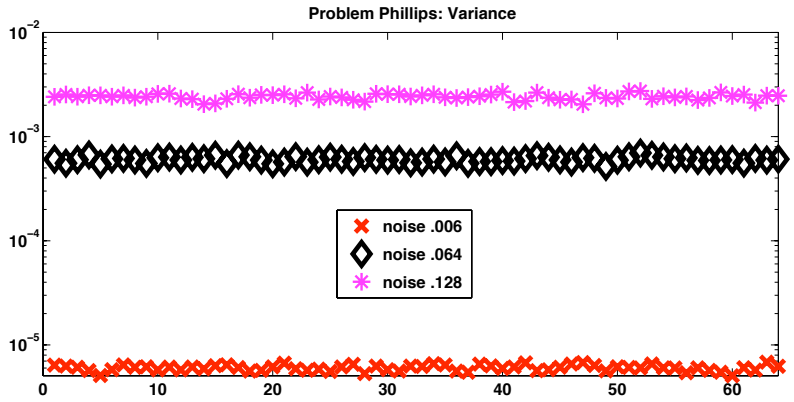



**Figure 3.** Illustration of the variance distribution of the noise in the right hand side for problem `phillips`, for the three noise levels used in the presented experiments. This is a plot of the variance $\sigma_{b_i}^2$ against index $i$ on the $x-$axis for each noise level.

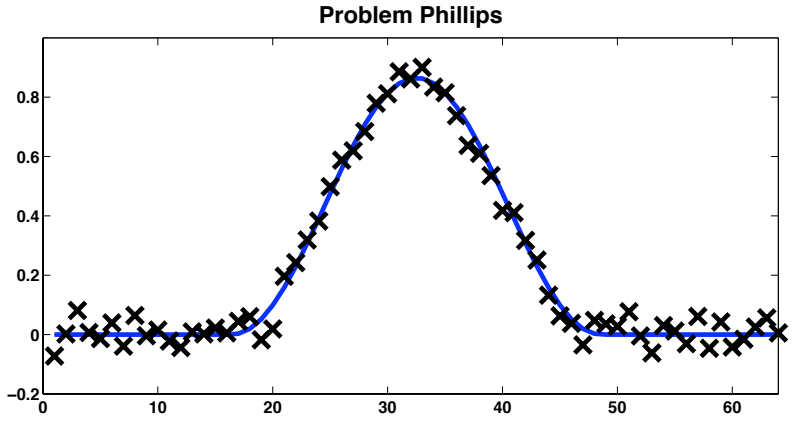



**Figure 4.** The reference solution $\mathbf{x}_0$ (crosses) used for problem `phillips`, as compared to the exact solution (solid line), in each case for the component of the vector against its index.

$\sigma$, the algorithm first proceeds by seeking to bracket the root. After this the Newton algorithm is implemented with a line search that maintains the current update of $\sigma$ within the current bracket, which is updated each step. Therefore the number of actual calculated $\sigma$ involved includes the bracketing step and the number of Newton steps. The number $K$ reported is the total number of calculated $\sigma$ in the Newton algorithm, including the bracketing step, see the Appendix. The error is the relative error in the solution $\|\mathbf{x} - \hat{\mathbf{x}}_{\rm rls}\|_2/\|\mathbf{x}\|_2$ and weighted predictive risk is $\|\tilde{A}(\hat{\mathbf{x}}_{\rm rls} - \mathbf{x})\|_2/m$. Average and standard deviation in values ($K$, $\sigma$, error and risk) are calculated over $500$ trials. The noise given is the average mean error resulting from using $\text{level} = .005, .05$ and $.1$ as noted above and is problem dependent, see Table 1. In all the experiments the algorithm is iterated to tolerance $|F| < .014$, which corresponds to high confidence $\alpha = .9999$ that the resulting $\sigma$ generates a functional which follows a $\chi^2$ distribution with $m + p - n$ degrees of freedom.

The data provided in Table 1 illustrate the robustness of the convergence of the Newton algorithm across noise levels and problem. The average number of required iterations is less than $10$ in all cases, and the standard deviation is always less than $4.5$.

| shaw | | phillips | | ilaplace | | heat | |
|---|---|---|---|---|---|---|---|
| noise | $K$ | noise | $K$ | noise | $K$ | noise | $K$ |
| 0.008 | 9.0(2.0) | 0.006 | 9.1(2.2) | 0.003 | 7.2(1.1) | 0.008 | 8.9(1.9) |
| 0.084 | 5.6(1.1) | 0.064 | 7.8(2.3) | 0.034 | 9.1(4.3) | 0.077 | 8.1(2.8) |
| 0.166 | 5.2(1.1) | 0.128 | 8.2(3.4) | 0.069 | 7.9(4.1) | 0.156 | 9.2(3.9) |
| 0.008 | 9.0(1.9) | 0.006 | 9.1(2.2) | 0.003 | 7.1(1.2) | 0.008 | 9.0(1.8) |
| 0.084 | 5.6(1.1) | 0.064 | 7.7(2.2) | 0.034 | 9.4(4.4) | 0.077 | 8.1(2.7) |
| 0.166 | 5.2(1.0) | 0.128 | 8.1(3.4) | 0.069 | 8.5(4.4) | 0.156 | 9.3(4.0) |

**Table 1.** Convergence characteristics of the Newton algorithm for the $\chi^2$-curve $n = 64$ over $500$ runs. Problem `shaw` uses the identity operator, and the other problems the first derivative operator. The first three rows are for white noise and the last three for colored noise. The data for $K$ are the mean and variance, in parentheses.

Tables 2-4 contrast the performance of the L-curve, GCV, UPRE and $\chi^2$-curve algorithms with respect to the relative error, risk, and regularization parameter, resp, for problems `shaw`, `phillips` and `ilaplace`, with noise for $\text{level} = .1$. With respect to the predictive risk and the relative error, the results of the $\chi^2$-curve are comparable to those with the UPRE statistical method. On the other hand, GCV, which is also statistically-based, is less robust, generally yielding larger error and risk, roughly comparable to results obtained with the L-curve. The results for the calculation of the regularization parameter show that the L-curve and GCV underestimate the regularization required, as compared to the UPRE and $\chi^2$. It can also be seen that the UPRE estimates of $\sigma_{\mathbf{x}}$ may be tighter than those achieved with the $\chi^2$-curve. We can interpret this in terms of the steepness, or lack of steepness, of the $\chi^2$ curve near its root. If the curve is very steep near the root, we would expect to obtain a very tight interval on the choice of $\sigma_{\mathbf{x}}$, ie for a small change in the number of degrees of freedom, equivalently of the width of the confidence interval, the change in $\sigma_{\mathbf{x}}$ is small. In contrast, if results for the $\chi^2$ curve are not tight across many trials, we can deduce that the general character of that problem leads to a curve which is not steep, and hence the solution is very dependent on the actual confidence interval imposed on the number of degrees of freedom in the solution. In such cases, we would need to make a much more severe tolerance in order to obtain less variance in the obtained values of $\sigma_{\mathbf{x}}$. ie. tighter results would be obtained by specifying a smaller tolerance. On the other hand, one cannot claim that any given method is totally ideal for picking a perfect regularization parameter. There will generally be a range of acceptable values for $\sigma_{\mathbf{x}}$, and its actual order of magnitude is much more significant for

generating reasonable results, than the specific obtained value. Indeed, there is always a trade-off between imposing tighter intervals and the computational cost.

| Problem | $cb$ | noise | L-Curve | GCV | UPRE | Chi |
|---|---|---|---|---|---|---|
| shaw | 2 | 0.166 | 0.1211(0.0266) | 0.4370(0.2934) | 0.1070(0.0593) | 0.1019(0.0235) |
| shaw | 3 | 0.166 | 0.1204(0.0262) | 0.4347(0.2919) | 0.1066(0.0579) | 0.1021(0.0202) |
| phillips | 2 | 0.128 | 0.1490(0.1191) | 0.1686(0.2018) | 0.1186(0.0884) | 0.1004(0.0010) |
| phillips | 3 | 0.128 | 0.1467(0.1099) | 0.1930(0.2370) | 0.1164(0.0810) | 0.1006(0.0014) |
| ilaplace | 2 | 0.069 | 0.3791(0.2186) | 0.2985(0.2464) | 0.1421(0.1068) | 0.1473(0.1122) |
| ilaplace | 3 | 0.069 | 0.3996(0.2374) | 0.2729(0.2357) | 0.1463(0.1178) | 0.1572(0.1364) |

**Table 2.** Mean and Standard Deviation of Error with $n = 64$ over 500 runs

| Problem | $cb$ | noise | L-Curve | GCV | UPRE | Chi |
|---|---|---|---|---|---|---|
| shaw | 2 | 0.166 | 0.0357(0.0088) | 0.0344(0.0127) | 0.0161(0.0082) | 0.0120(0.0036) |
| shaw | 3 | 0.166 | 0.0354(0.0089) | 0.0342(0.0126) | 0.0162(0.0081) | 0.0125(0.0038) |
| phillips | 2 | 0.128 | 0.0379(0.0106) | 0.0268(0.0115) | 0.0298(0.0112) | 0.0225(0.0063) |
| phillips | 3 | 0.128 | 0.0379(0.0107) | 0.0283(0.0128) | 0.0297(0.0111) | 0.0229(0.0064) |
| ilaplace | 2 | 0.069 | 0.0367(0.0081) | 0.0244(0.0137) | 0.0194(0.0103) | 0.0169(0.0071) |
| ilaplace | 3 | 0.069 | 0.0373(0.0086) | 0.0217(0.0124) | 0.0198(0.0105) | 0.0172(0.0079) |

**Table 3.** Mean and Standard Deviation of Risk with $n = 64$ over 500 runs

| Problem | $cb$ | noise | L-Curve | GCV | UPRE | Chi |
|---|---|---|---|---|---|---|
| shaw | 2 | 0.166 | 0.6097(0.3993) | 6.1070(5.4358) | 0.1219(0.4446) | 0.1683(0.6863) |
| shaw | 3 | 0.166 | 0.6043(0.3989) | 6.6179(8.7190) | 0.1181(0.4235) | 0.1720(0.3755) |
| phillips | 2 | 0.128 | 0.0902(0.1958) | 0.0810(0.1922) | 0.0283(0.0880) | 0.0061(0.0089) |
| phillips | 3 | 0.128 | 0.0860(0.1801) | 0.1132(0.2589) | 0.0260(0.0799) | 0.0065(0.0116) |
| ilaplace | 2 | 0.069 | 0.1354(0.1478) | 0.5316(1.5602) | 0.0207(0.0315) | 0.0421(0.1018) |
| ilaplace | 3 | 0.069 | 0.1682(0.2452) | 0.2969(1.0878) | 0.0218(0.0338) | 0.0456(0.1421) |

**Table 4.** Mean and Standard Deviation of Sigma with $n = 64$ over 500 runs

### *5.3. Estimating data error: Example from Hydrology*

In this section we use the $\chi^2$-curve method to estimate soil hydraulic propertiesa, and note that in this application it used to find the weight $\sigma_{\mathbf{b}}$ on field measurements $\mathbf{b}$, as described in Section 2.2. The initial parameter estimates $\mathbf{x}_0$ and their respective covariance $C_{\mathbf{x}} = \mathrm{diag}(\sigma_{x_i}^2)$ are based on laboratory measurements, and specified *a priori*. In this example, the parameter

estimates obtained by the $\chi^2$-curve method optimally combine field and laboratory data, and give an estimate of field measurement error, i.e. $\sigma_{\mathbf{b}}$.

In nature, the coupling of terrestrial and atmospheric systems happens through soil moisture. Water must pass through soil on its way to groundwater and streams, and soil moisture feeds back to the atmosphere via evapotranspiration. Consequently, methods to quantify the movement of water through unsaturated soils are essential at all hydrologic scales. Traditionally, soil moisture movement is simulated using Richards' [19] equation for unsaturated flow:

$$\frac{\partial \theta}{\partial t} = \nabla \cdot (K \nabla h) + \frac{\partial K}{\partial z}. \tag{36}$$

$\theta$ is the volumetric moisture content, $h$ pressure head, and $K(h)$ hydraulic conductivity. Solution of Richards' equation requires reliable inputs for $\theta(h)$ and $K(h)$. These relationships are collectively referred to as soil-moisture characteristic curves, and they are typically highly nonlinear functions. These curves are commonly parameterized using the van Genuchten [22] and Mualem [16] relationships:

$$\begin{aligned} \theta(h) &= \theta_r + \frac{(\theta_s - \theta_r)}{(1 + |\alpha h|^n)^m} \qquad h < 0 \\ K(h) &= K_s \theta_e^l \left(1 - (1 - \theta_e^{1/m})^m\right)^2 \qquad h < 0 \\ \theta_e &= \frac{\theta(h) - \theta_r}{\theta_s - \theta_r}. \end{aligned} \tag{37}$$

These equations contain seven independent parameters: $\theta_r$ and $\theta_s$ are the residual and saturated water contents (both with units $\text{cm}^3\text{cm}^{-3}$), respectively, $\alpha$ ($\text{cm}^{-1}$), $n$ and $m$ (commonly set $= 1 - 1/n$) are unit-less empirical fitting parameters, $K_s$ is the saturated hydraulic conductivity (cm/sec) and $l$ is the pore connectivity parameter (-). (Note that the variables used in this application, $K$, $n$, $m$, $\alpha$, $r$, $s$, are common notation for the parameters in the model, and do not represent the number of iterations, size of the problem, confidence interval etc. as in previous sections.) We focus on parameter estimates for $\theta(h)$ to be used in Richards' equation.

Field studies and laboratory measurements of soil moisture content and pressure head in the Dry Creek catchment near Boise, Idaho are used to obtain $\theta(h)$ estimates. This watershed is typical of small watersheds in the Idaho Batholith and hydrologic studies have been conducted there since 1998 under grants from the NASA Land Surface Hydrology program and the USDA National Research Initiative, [14]. Detailed hill slope and small-catchment studies have been ongoing in two locations at low and intermediate elevations. Measurements used here to test the $\chi^2$-curve method are from the intermediate elevation small catchment which drains $0.02\ km^2$. The north facing slope is currently instrumented with a weather station, two soil pits recording temperature and moisture content at four depths, and four additional soil pits recording moisture content and pressure head with TDR/tensiometer

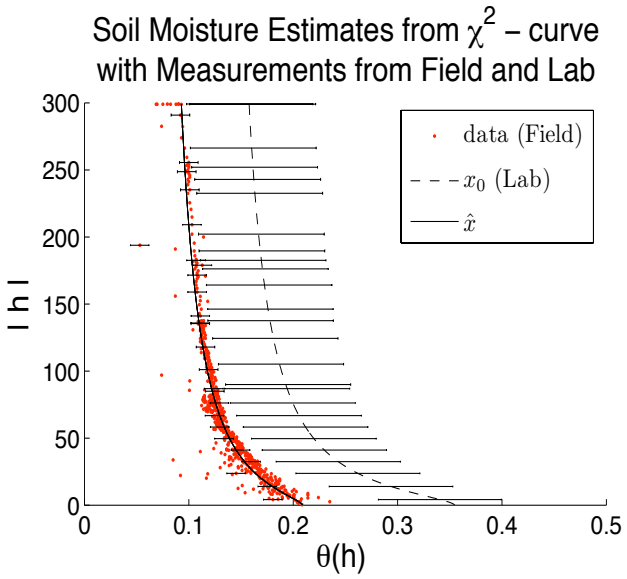



**Figure 5.** Soil water retention curves, $\theta(h)$, observed in the field, and in the laboratory. The laboratory curve gives initial parameter estimates, while field measurements are the "data", $\mathbf{b}$.

pairs. We will show soil water retention curves from one of these pits.

Laboratory measurements were made on undisturbed samples (approximately 54 mm in diameter and 30 mm long) taken from the field soil pits at the same depths, but up to 50 cm away from the location of the in-situ measurements so as to not influence subsequent measurements. These samples were subjected to multi-step outflow tests in the Laboratory [4]. The curves that fit the laboratory data do not necessarily reflect what is observed in the field, see Figure 5. The goal is to obtain parameter estimates $\mathbf{x}$ which contain soil moisture and pressure head information from both laboratory and field measurements, within their uncertainty ranges. We rely on the laboratory measurements for good first estimates of the parameters $\mathbf{x}_0 = [\theta_r, \theta_s, \alpha, n, m]$ and their standard deviations $\sigma_{x_i}$. It takes 2-3 weeks to obtain one set of laboratory measurements, but this procedure is done multiple times from which we obtain standard deviation estimates and form $C_{\mathbf{x}} = \mathrm{diag}(\sigma^2_{\theta_r}, \sigma^2_{\theta_s}, \sigma^2_{\alpha}, \sigma^2_{n}, \sigma^2_{m})$. These standard deviations account for measurement technique or error. however, measurements on this core may not accurately reflect soils in entire watershed region.

In order to include what is observed in the field, the initial parameter estimates from the laboratory, $\mathbf{x}_0$, are updated with measurements collected in-situ. However, we also cannot entirely rely on the in-situ data because it contains error due to incomplete saturation, spatial variability, measurement technique and error. This means we must also specify a data error weight $C_{\mathbf{b}}$ *a priori*. It is not possible to obtain repeated measurements of field data, and get uncertainty estimates as was done with laboratory cores. We instead estimate $C_{\mathbf{b}}$ by using the $\chi^2$-curve method. This requires a linear model, and we use the following technique to

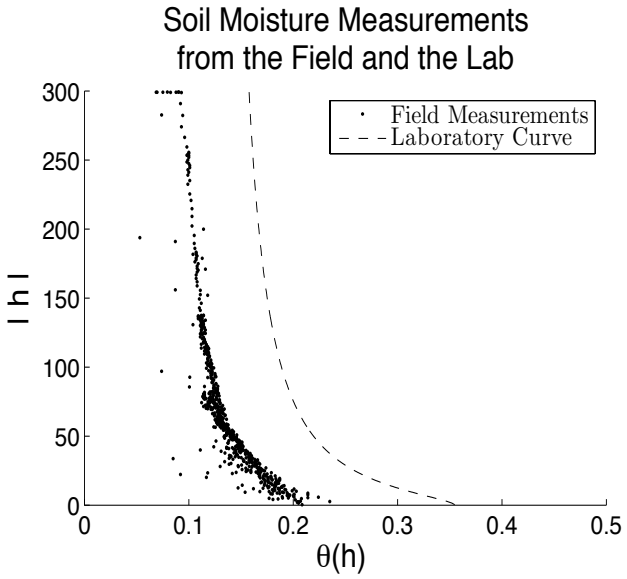



**Figure 6.** The curve resulting from the value of $\hat{\mathbf{x}}$ found by the $\chi^2$-curve method optimally combines field and laboratory measurements, within their standard deviation ranges.

simulate the nonlinear behavior of (37). Let

$$\theta(h, \mathbf{x}) \approx \theta(h, \mathbf{x}_0) + \left.\frac{\partial \theta}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_0) \tag{38}$$

$$= A(h, \mathbf{x}_0)\mathbf{x} + \mathbf{q}(h, \mathbf{x}_0), \tag{39}$$

where

$$A(h, \mathbf{x}_0) = \left[ \frac{\partial \theta(h, \mathbf{x}_0)}{\partial \theta_r} \; \frac{\partial \theta(h, \mathbf{x}_0)}{\partial \theta_s} \; \frac{\partial \theta(h, \mathbf{x}_0)}{\partial \alpha} \; \frac{\partial \theta(h, \mathbf{x}_0)}{\partial n} \; \frac{\partial \theta(h, \mathbf{x}_0)}{\partial m} \right],$$

and

$$\mathbf{q}(h, \mathbf{x}_0) = \theta(h, \mathbf{x}_0) - A(h, \mathbf{x}_0)\mathbf{x}_0.$$

Then $b_i = \theta(h_i, \mathbf{x}) - \mathbf{q}(h_i, \mathbf{x}_0)$ has dimension equal to the number of measurements $N$, while $A$ has dimension $N \times 5$. An optimal $\hat{\mathbf{x}}$ is found by the $\chi^2$-curve method, $\mathbf{x}_0$ is updated with it, and (39) is iterated. The results are shown in Figure 6 where we plot the initial estimate $\mathbf{x}_0$ and field measurements with their standard deviations, along with the final curve found by iterating the $\chi^2$-curve method. The final estimate $\hat{\mathbf{x}}$ converged in three iterations on the linear model (39) . The standard deviation for $\mathbf{x}_0$ was specified *a priori* while the standard deviation for the data was calculated with the $\chi^2$-curve method, and is estimated to be $0.02775$. We note that the large standard deviation on $\mathbf{x}_0$ observed in the laboratory, resulted in optimal estimates $\hat{\mathbf{x}}$ which more closely resemble what was observed in the field. The value of the $\chi^2$ variable in this example exactly matched the number of data: 694.

## 6. Conclusions

A cost effective and robust algorithm for finding the regularization parameter for the solution of regularized least squares problems has been presented, when *a priori* information on either model or data error is available. The algorithm offers a significant advantage as compared to other statistically-based approaches because it determines the unique root, when the root exists, of a nonlinear monotonic decreasing scalar function. Consistent with the quadratic convergence properties of Newton's method, the algorithm converges very quickly, in, on average, no more than ten steps. Compared with the UPRE or GCV approaches this offers a considerable cost advantage, and avoids the problem with the multiple minima of the GCV and UPRE. In accord with the comparison of statistically-based methods discussed in [23], we conclude that a statistically-based method should be used whenever such information is available. While we have not yet compared with new work described in [20] which is statistically-based but appears to be more computationally demanding, our results suggest that our new Newton $\chi^2$-curve method is competitive. These positive results are encouraging for the extension of the method, as suggested in [15], for a more general covariance structure of the data error. This will be the subject of future research. Moreover, we note that the theory presented here is only valid as given when $\mathbf{x}_0$ is the appropriate expected mean of the model parameters. Correspondingly the results presented were only for the case in which an estimate of $\mathbf{x}_0$ is provided. The extension and validation of the method when $\mathbf{x}_0$ is not available is also a subject of future research.

## Acknowledgements

Statistics Professor Randall Eubank, Arizona State University, School of Mathematics and Statistics, assisted the authors by verification of the quoted statistical theory presented in Theorem 3.1. Professor Jim McNamara, Boise State University, Department of Geosciences and Professor Molly Gribb, Boise State University, Department of Civil Engineering supplied the field and laboratory data, respectively, for the Hydrological example.

## Appendix

The $\chi^2$-curve method uses a basic Newton iteration with an initial bracketing step and a line search to maintain the new value of $\sigma$ within the given bracket. The generic algorithm in all cases is the same, and just depends on the functional for calculating $F(\sigma)$, dependent on either the SVD (14-16) or GSVD (29) as appropriate, and whether the goal is to find $\sigma_{\mathbf{b}}$ or $\sigma_{\mathbf{x}}$. Given the optimal $\sigma$ solution $\hat{\mathbf{x}}$ can be obtained using (18) or (30).

**Algorithm 2 (Find $\sigma$ which is a root of $F(\sigma)$)** *Given functionals for evaluation of $F(\sigma)$ and $\mathbf{step}(\sigma) = F(\sigma)/(\sigma F'(\sigma))$, tolerance $\tau$, maximum number of steps $K_{\max}$, initial residual $\mathbf{r} = \mathbf{b} - A\mathbf{x}_0$, and the degrees of freedom $m-n+p$. Set $\sigma = 1$, $k = 1$, $\alpha_0 = 1$, and calculate $F(\sigma)$.*

**Bracket the root:**
**Find** $\sigma_{\min} > 0$ and $\sigma_{\max} < 10^6$ such that $F(\sigma_{\min}) > 0 > F(\sigma_{\max})$
Use Logarithmic search on $\sigma$ and increment $k$ for each $F(\sigma)$.
**Stop** if no bracket exists.
**Newton updates with line search:**
**While** $F(\sigma) > \tau$ & $k < K_{\max}$ **Do**
Set $\alpha = \alpha_0$. Evaluate $\mathbf{step}(\sigma)$
**Until** $\sigma_{\text{new}}$ within bracket **Do** $\sigma_{\text{new}} = \sigma(1 + \alpha\,\mathbf{step})$, $\alpha = \alpha/2$.
Set $\sigma = \sigma_{\text{new}}$, $k = k + 1$, update $F(\sigma)$

As previously noted in Sections 2.2, 3.1 the algorithms can be modified to handle the case that $C_{\mathbf{x}}$ is given and $C_{\mathbf{b}}$ is to be found. For the case when $D = I$ we use from (21)

$$\begin{aligned} F(\sigma) &= \tilde{\mathbf{s}}^T\mathbf{s} - m, \quad \tilde{s}_i = \frac{s_i}{\sigma_i^2 + \sigma^2}, \quad \sigma_i = 0, i > r \\ F'(\sigma) &= -2\sigma\|\tilde{\mathbf{s}}\|^2 \\ \hat{\mathbf{x}} &= \mathbf{x}_0 + L_{\mathbf{x}}V\mathbf{t}, \quad t_i = \frac{\sigma_i}{\sigma^2 + \sigma_i^2} \quad \text{and} \\ \tilde{C}_{\mathbf{x}} &= \sigma^2 L_{\mathbf{x}} V \operatorname{diag}\left(\frac{1}{\sigma_i^2 + \sigma^2}\right) V^T L_{\mathbf{x}}^T. \end{aligned}$$

Otherwise, for the GSVD case the equivalent results are obtained with $\mathbf{s} = U^T\mathbf{r}$, $\gamma_i = 0, i > p$, and $\tilde{m} = m - n + p$.

$$\begin{aligned} F(\sigma) &= \mathbf{s}^T\tilde{\mathbf{s}} - \tilde{m} \quad \tilde{s}_i = \frac{s_i}{\sigma^2 + \gamma_i^2}, i = 1 \ldots m, \gamma_i \neq 0, \tilde{s}_i = 0, \text{ otherwise}, \\ F'(\sigma) &= -2\sigma\|\tilde{\mathbf{s}}\|^2 \\ t_i &= \frac{\gamma_i \tilde{s}_i}{\mu_i}, i = 1 \ldots p, \quad t_i = s_i, i = p+1 \ldots n, \quad \text{and } t_i = 0 \text{ otherwise} \\ \hat{\mathbf{x}} &= \mathbf{x}_0 + (X^T)^{-1}\mathbf{t}, \\ \tilde{C}_{\mathbf{x}} &= \sigma^2 (X^T)^{-1} \left(\operatorname{diag}\left(\frac{1}{\nu_i^2 + \sigma^2\mu_i^2}\right), I_{n-p}\right) X^{-1}. \end{aligned}$$

## Bibliography


[1] Bennett, A., 2005 *Inverse Modeling of the Ocean and Atmosphere* (Cambridge University Press) p 234.

[2] Chung, J., Nagy, J., and O'Leary, D. P., 2008, *A Weighted GCV Method for Lanczos Hybrid Regularization*, ETNA, Vol. 28, 149-167.

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

[4] Figueras, J., and Gribb, M. M., *A new user-friendly automated multi-step outflow test apparatus*, Vadose Zone Journal, (submitted).

[5] Golub, G. H. and van Loan, C., 1996, *Matrix Computations*, John Hopkins Press, Baltimore, 3rd ed..

[6] Hansen, P. C., 1989, Regularization, GSVD and Truncated GSVD, *BIT*, **6**, 491-504.

[7] Hansen, P. C., 1994, *Regularization Tools: A Matlab Package for Analysis and Solution of Discrete Ill-posed Problems*, Numerical Algorithms, **6** 1-35.

[8] Hansen, P. C., 1998 *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion* (SIAM) Monographs on Mathematical Modeling and Computation **4**.

[9] Hansen, P. C., Nagy, J. G., and O'Leary, D. P., 2006, *Deblurring Images, Matrices, Spectra and Filtering*, (SIAM), Fundamentals of Algorithms.

[10] Kilmer, M. E. and O'Leary, D. P., 2001, *Choosing regularization parameters in iterative methods for ill-posed problems*, SIAM J. Numer. Anal. Appl. **22**, 1204-1221.

[11] Kilmer, M. E., Hansen, P. C., and Español, M. I., 2007 *A Projection-based approach to general-form Tikhonov regularization*, SIAM J. Sci. Comput., **29**, 1, 315–330.

[12] Marquardt, D. W., 1970, *Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation*, Technometrics, **12**, 3, 591 - 612.

[13] MATLAB is a registered mark of MathWorks, Inc., MathWorks Web Site, `http://.mathworks.com`.

[14] McNamara, J. P., Chandler, D. G., Seyfried, M., and Achet, S., 2005. *Soil moisture states, lateral flow, and streamflow generation in a semi-arid, snowmelt-driven catchment*, Hydrological Processes, **19**, 4023-4038.

[15] Mead, J., 2008, *Parameter estimation: A new approach to weighting a priori information*, J. Inv. Ill-posed Problems, **16**, 2.

[16] Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* , **12** 513-.

[17] O'Leary, D. P. and Simmons, J. A., 1989, it A Bidiagonalization-Regularization Procedure for Large Scale Discretizations of Ill-Posed Problems, SIAM Journal on Scientific and Statistical Computing, **2**, 4, 474-489.

[18] Rao, C. R., 1973, *Linear Statistical Inference and its applications*, Wiley, New York.

[19] Richards, L.A., Capillary conduction of liquids in porous media, *Physics 1*, 318-333.

[20] Rust, B. W., and O'Leary, D. P., 2008, Residual periodograms for choosing regularization parameters for ill-posed problems, Inverse Problems **24** 034005.

[21] Tarantola, A., 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM).

[22] van Genuchten, M. Th., 1980, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* , **44**, 892-898.

[23] Vogel, C. R., 2002, *Computational Methods for Inverse Problems*, (SIAM), Frontiers in Applied Mathematics.