This is a repository copy of *The application of regularisation to variable selection in statistical modelling*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/179180/

Version: Accepted Version

Article:

Winkler, J. orcid.org/0000-0002-4629-8928, Mitrouli, M. and Koukouvinos, C. (2022) The application of regularisation to variable selection in statistical modelling. Journal of Computational and Applied Mathematics, 404. 113884. ISSN 0377-0427

https://doi.org/10.1016/j.cam.2021.113884

Article available under the terms of the CC-BY-NC-ND licence (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

The application of regularisation to variable selection in statistical modelling

Joab R. Winkler,[a] Marilena Mitrouli,[b] Christos Koukouvinos[c]

[a]*Department of Computer Science, The University of Sheffield, Regent Court, 211 Portobello, Sheffield S1 4DP, United Kingdom*

[b]*Department of Mathematics, National and Kapodistrian University of Athens, Panepistimioupolis, GR157 84, Athens, Greece*

[c]*Department of Mathematics, National Technical University of Athens, Zografou 15773, Athens, Greece*

j.r.winkler@sheffield.ac.uk, mmitroul@math.uoa.gr, ckoukouv@math.ntua.gr

Abstract

The aim of variable selection is the identification of the most important predictors that define the response of a linear system. Many techniques for variable selection use a constrained least squares (LS) formulation in which the constraint is imposed in the 1-norm (the lasso), or the 2-norm (Tikhonov regularisation), or a linear combination of these norms (the elastic net). It is always assumed that a constraint must necessarily be imposed, but the consequences of its imposition have not been addressed. This assumption is considered in this paper and it is shown that the correct application of Tikhonov regularisation to the overdetermined LS problem min $\|Ax-b\|_2$ requires that A and b satisfy a condition $\mathcal{C}$. If this condition is satisfied, then the solution of the LS problem with this constraint is numerically stable and the regularisation error e between the solution of this problem and the solution of the LS problem is small. If, however, the condition $\mathcal{C}$ is not satisfied, then the error e is large. The condition $\mathcal{C}$ is derived from a refined normwise condition number of the solution of the LS problem. The paper includes examples of regularisation and variable selection with correlated variables that illustrate the theory in the paper.

Key words: Ridge regression (Tikhonov regularisation); condition number; discrete Picard condition; regularisation error; lasso

1 Introduction

The identification of the most important parameters (predictors) that define the response of a linear system is important for statistical modelling of physical systems. Many predictors are included in the initial model of the linear system, but the identification of the most important predictors requires a reduction in the number of predictors from the initial large number by the elimination of some of them, such that a simpler model is obtained. This process of the elimination of predictors that are not important for the characterisation of the linear system is called variable selection and it arises in several applications, including text mining [2] and medical imaging [12].

A linear model of a system is defined by the equation

$$y = X\beta + \epsilon, \tag{1}$$

where $X \in \mathbb{R}^{n\times p}$, $n \geq p$, $\operatorname{rank} X = p$, $y \in \mathbb{R}^n$, $\beta \in \mathbb{R}^p$ and ϵ is the noise. Equation (1) is solved in the least squares (LS) sense for β, which yields the equation

$$X^T X\beta = X^T y, \tag{2}$$

whose solution is

$$\beta_{\text{LS}} = X^\dagger y, \qquad X^\dagger = \left(X^T X\right)^{-1} X^T. \tag{3}$$

The solution β_{LS} is, in general, not sparse and it is therefore not suitable for variable selection. Also, a small relative change in y may cause a large relative change in β_{LS}, which is unsatisfactory. These problems are addressed by the imposition of a constraint on β_{LS}, such that a modified solution $\hat{\beta}(\lambda_1, \lambda_2)$ of (1) is sought,

$$\hat{\beta}(\lambda_1, \lambda_2) = \arg\min_{\beta\in\mathbb{R}^p} \left\{ \|y - X\beta\|_2^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 \right\}, \tag{4}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are parameters whose values define the severity with which the constraints are imposed. The condition $\lambda_1 = 0$, $\lambda_2 > 0$ yields Tikhonov regularisation [6], the condition $\lambda_1 > 0$, $\lambda_2 = 0$ yields the lasso [10], and the condition $\lambda_1, \lambda_2 > 0$ yields the elastic net [14]. Tikhonov regularisation is called ridge regression in the machine learning and statistics literature, but the term Tikhonov regularisation is used in this paper.

The function $\|v\|_0$ is a measure of the sparsity of v because it is equal to the number of its non-zero elements, but it is not a norm because it does not satisfy the triangle inequality. The best convex approximation of $\|v\|_0$ is $\|v\|_1$, and thus the condition $\lambda_1 > 0, \lambda_2 = 0$ is used to impose sparsity on β_{LS}. The condition $\lambda_1 = 0, \lambda_2 > 0$ is different because it reduces $\|\beta_{\text{LS}}\|$ but it does not impose sparsity on β_{LS}.

Tikhonov regularisation is discussed extensively in the numerical analysis literature [6] but this literature is not cited by researchers in computational statistics and thus concepts from computational linear algebra, for example, regularisation error and condition estimation, are not considered in computational statistics. It is shown in this paper that Tikhonov regularisation can be applied to (2) only if X and y satisfy the discrete Picard condition [6, §4.5], and more generally, a regularised solution must satisfy two fundamental properties when λ_1 and λ_2 assume their optimal values, λ_1^* and λ_2^* respectively:

(1) It is numerically stable, that is, a small relative error in y causes a small relative error in $\hat{\beta}(\lambda_1^*, \lambda_2^*)$.
(2) The error $e(\lambda_1^*, \lambda_2^*)$ between β_{LS} and $\hat{\beta}(\lambda_1^*, \lambda_2^*)$, that is, the regularisation error, is small,

$$e(\lambda_1^*, \lambda_2^*) = \frac{\left\|\hat{\beta}(\lambda_1^*, \lambda_2^*) - \beta_{\text{LS}}\right\|}{\|\beta_{\text{LS}}\|} \ll 1.$$

Consider the situation defined by $\lambda_1 = 0$ and $\lambda_2 > 0$, in which case the singular value decomposition (SVD) of X allows considerable analytical progress to be made. Closed form expressions for the error $e(0, \lambda_2)$ and the stability of $\hat{\beta}(0, \lambda_2)$ are derived, and it is shown that the requirement of a stable solution and a small error are satisfied if X and y satisfy the discrete Picard condition and λ_2 assumes its optimal value λ_2^*. These issues are considered in detail in this paper and examples that illustrate the theoretical results are included.

The algorithm for variable selection is considered in Section 2 and it is shown that it contains two computations, called the forward and inverse problems. Expressions for their effective condition numbers, which are refined measures of their stability, are derived in Section 3, and they allow detailed analysis of Tikhonov regularisation, which is considered in Section 4. The discrete Picard condition is introduced and it is shown that the satisfaction of this condition is critical for the correct application of Tikhonov regularisation because it guarantees that the solution $\hat{\beta}(0, \lambda_2^*)$ satisfies the two properties stated above.

The matrix X is arbitrary in Sections 3 and 4, but it is restricted to a design matrix with correlated covariates in Section 5. Examples of variable selection with correlated covariates and the application of regularisation to (2) are in Section 6, and the paper is summarised in Section 7.

The work described in this paper extends the work in [13] because several important issues are addressed:

(1) Only the effective condition number of the inverse problem is considered in [13], but this paper considers the effective condition numbers of the forward and inverse problems. This allows analysis of the uncertainty principle, which is discussed briefly in [13], and it is shown that the form of the uncertainty principle is dependent on the order in which the forward and inverse problems are solved. An example of regression that shows the importance of the uncertainty principle is in [13, §3].
(2) Numerical issues associated with the computation of the effective condition numbers of the forward and inverse problems are considered in Section 4, but they are not addressed in [13]. It is shown that a distinction must be made between the stability/instability of the forward and inverse problems, and the stability/instability of their effective condition numbers.
(3) The discrete Picard condition is associated with the solution of an inverse problem, but the literature does not show that it can also be associated with the solution of the forward problem in variable selection. This association arises because variable selection requires that both the forward and inverse problems be solved, and it is shown that this condition allows theoretical and numerical issues associated with these problems to be considered.
(4) Variable selection with correlated covariates is discussed in Section 5, but this issue is not considered in [13]. The examples in Section 6 show the different forms of the graphs of two methods, the generalised cross validation (GCV) and L-curve, for the calculation of the optimal value of the regularisation parameter. In particular, these graphs assume radically different forms, dependent upon the need, or otherwise, to impose Tikhonov regularisation. This result emphasizes the care with which regularisation must be applied because its application when it is not required causes a degradation in the computed solution. This distinction between the correct and incorrect application of Tikhonov regularisation is not addressed in the statistics literature.

2 Variable selection

The algorithm for variable selection is shown in Algorithm 1 and it is assumed that $\lambda_1 = 0$ and $\lambda_2 \neq 0$, and the algorithm for the more general condition $\lambda_1, \lambda_2 \neq 0$ follows easily. The value of λ_2 is defined in lines 13 and 14 of the algorithm by the GCV [1], [3], [4, Chapter 15] and [6, §7.4], and by the L-curve [4, Chapter 15], [5] and [6, §4.6]. The two fundamental computations are in lines 9 and 10 of the algorithm, and extensive reference is made to them. The best solutions from Algorithm 1 are obtained by selecting the variables associated with the columns of X that correspond to the dominant non-zero entries in the solution vector β. This selection is shown in Example 6.3.

Algorithm 1 Variable selection

Input: X, β, N % N is the number of trials

Output: Errors e_0, e_1 and e_2 of, respectively, the LS solution, and the regularised solutions using the GCV and the L-curve to determine the regularisation parameter.

% Initialise the scalars that store the errors.
$e_0 \leftarrow 0,\ e_1 \leftarrow 0,\ e_2 \leftarrow 0$
% Start the loop for the number of trials N in the presence of noise.
for $i = 1 : N$ **do**
  % Compute the response in the presence of noise ϵ and solve
  % the LS problem.
  $y \leftarrow X\beta + \epsilon$
  $\beta_{\text{LS}} \leftarrow X^{\dagger}y$
  % If regularisation is required, compute the regularisation parameters
  % using the GCV and L-curve, and apply Tikhonov regularisation.
  $\lambda^{\text{GCV}} = \text{GCV}(X, y)$
  $\lambda^{\text{Lcurve}} = \text{Lcurve}(X, y)$
  $\beta\left(\lambda^{\text{GCV}}\right) = \text{Tikhonov}\left(X, y, \lambda^{\text{GCV}}\right)$
  $\beta\left(\lambda^{\text{Lcurve}}\right) = \text{Tikhonov}\left(X, y, \lambda^{\text{Lcurve}}\right)$
  % Calculate the errors in β_{LS}, $\beta\left(\lambda^{\text{GCV}}\right)$ and $\beta\left(\lambda^{\text{Lcurve}}\right)$.
  $e_0 \leftarrow e_0 + \frac{\|\beta - \beta_{\text{LS}}\|_2}{\|\beta\|_2}$
  $e_1 \leftarrow e_1 + \frac{\left\|\beta - \beta\left(\lambda^{\text{GCV}}\right)\right\|_2}{\|\beta\|_2}$
  $e_2 \leftarrow e_2 + \frac{\left\|\beta - \beta\left(\lambda^{\text{Lcurve}}\right)\right\|_2}{\|\beta\|_2}$
end for
% Calculate the mean errors.
$e_0 \leftarrow \frac{e_0}{N}, \quad e_1 \leftarrow \frac{e_1}{N}, \quad e_2 \leftarrow \frac{e_2}{N}$

Proposition 2.1 Consider (1) in which y is computed from X, β and ϵ. This model leads to the linear system

$$y = X(\beta + \delta\beta),$$

where $\delta\beta$ is the component of ϵ that lies in the column space $\mathcal{R}(X)$ of X.

Proof The noise ϵ is written as $\epsilon = \epsilon_1 + \epsilon_2$ where $\epsilon_1 \in \mathcal{R}(X)$ and ϵ_2 is orthogonal to $\mathcal{R}(X)$. There therefore exists a vector $\delta\beta$ such that $\epsilon_1 = X\delta\beta$, and thus

$$\epsilon = X\delta\beta + \epsilon_2 \qquad \text{and} \qquad \epsilon_1^T\epsilon_2 = \delta\beta^T X^T \epsilon_2 = 0,$$

from which it follows that

$$y = X\beta + \epsilon = X(\beta + \delta\beta) + \epsilon_2.$$

If y_1 and y_2 are defined as

$$y_1 = X(\beta + \delta\beta) \qquad \text{and} \qquad y_2 = \epsilon_2,$$

then y_1 lies in $\mathcal{R}(X)$, and y_2 is composed of noise and is orthogonal to $\mathcal{R}(X)$. It follows that only y_1, and not y_2, contains information included in X. $\square$

Proposition 2.1 leads to the following definitions of the forward and inverse problems.

Definition 2.1 (The forward problem) Given a matrix $X \in \mathbb{R}^{n \times p}$ and a vector $\beta \in \mathbb{R}^p$, the forward problem is defined as the computation of y,

$$y := X\beta. \tag{5}$$

Definition 2.2 (The inverse problem) Given a matrix $X \in \mathbb{R}^{n \times p}$, $n \geq p$, of full column rank p and a vector $y \in \mathbb{R}^n$, the inverse problem is defined as the computation of β_{LS},

$$\beta_{\text{LS}} := X^{\dagger} y. \tag{6}$$

Lines 9 and 10 in Algorithm 1 show that the response y of the forward problem is the input to the inverse problem, and these two problems should therefore be considered together. There are two issues that must be addressed:

(1) If y is sensitive to a small relative error $\Delta\beta$ in β, then it is computationally unreliable and the consequences on β_{LS} must be considered.
(2) If y is stable with respect to $\Delta\beta$, then $\Delta y \approx \Delta\beta$, but it does not follow that the relative error $\Delta\beta_{\text{LS}}$ in β_{LS} satisfies $\Delta\beta_{\text{LS}} \approx \Delta y$. The satisfaction of these approximations is desirable because it guarantees that the forward and inverse problems are well conditioned, but the relationship between X, β and y must be considered such that the conditions when these approximations are satisfied, and are not satisfied, are defined.

These issues are important because it is shown in Section 4 that if regularisation is applied to the inverse problem when it is numerically stable, then the regularisation error in the regularised solution is large. This result requires that the numerical condition of the forward and inverse problems be considered, and this issue is addressed in Section 3.

3 The numerical condition of the forward and inverse problems

It was shown in Section 2 that the solution of the forward problem defines the input to the inverse problem in variable selection. These problems are therefore related, but it is convenient to consider them separately and then combine the results, such that the two points raised in Section 2 are addressed.

Section 3.1 considers the condition number $\kappa(X) = \|X\| \left\|X^{\dagger}\right\|$ of X, and refined condition numbers, called effective condition numbers, of the forward and inverse problems, and their advantages with respect to $\kappa(X)$ for condition estimation, are discussed. It is shown in Section 3.2 that the product of the effective condition numbers of these problems in variable selection is equal to $\kappa(X)$ because the forward problem is solved before the inverse problem, and the implications of this result are discussed. [1]

3.1 Condition numbers and effective condition numbers

The condition number $\kappa(X)$ of X is not a good measure of the numerical condition of the forward and inverse problems because it is a function of X, and it is independent of β for the forward problem and independent of y for the inverse problem, but the solution y of the forward problem is a function of X and β, and the solution β_{LS} of the inverse problem is a function of X and y. It is therefore necessary to derive refined expressions for the numerical condition of the forward and inverse problems. The effective condition number of the inverse problem is introduced in [13, §4], and the analysis is extended in this paper because it allows an uncertainty principle to be considered in significantly greater detail than in [13, §6].

Let the relative errors $\Delta\beta$, $\Delta\beta_{\text{LS}}$ and Δy in β, β_{LS} and y be, respectively,

$$\Delta\beta = \frac{\|\delta\beta\|}{\|\beta\|}, \qquad \Delta\beta_{\text{LS}} = \frac{\|\delta\beta_{\text{LS}}\|}{\|\beta_{\text{LS}}\|} \qquad \text{and} \qquad \Delta y = \frac{\|\delta y\|}{\|y\|}. \tag{7}$$

Expressions for the condition numbers and effective condition numbers of the forward and inverse problems are derived in Theorems 3.1, 3.2 and 3.3, and the difference between the condition numbers and the effective condition numbers is evident. In particular, the effective condition number of the forward problem is equal to the maximum value of $\gamma = {}^{\Delta y}/_{\Delta\beta}$ for all vectors $\delta\beta$ and the given vector β, and the condition number $\kappa(X)$ is equal to the maximum value of γ

[1] A slightly different result is obtained if the inverse problem is solved before the forward problem.

for all vectors $\delta\beta$ and all vectors β. It is shown that the equivalent result for the inverse problem is more complex.

Theorem 3.1 The effective condition number $\eta^{\text{fwd}}(X,\beta)$ and the condition number $\kappa^{\text{fwd}}(X)$ of the forward problem (5) are, respectively,

$$\eta^{\text{fwd}}(X,\beta) = \max_{\delta\beta\in\mathbb{R}^p} \frac{\Delta y}{\Delta\beta} = \frac{\|X\|\,\|\beta\|}{\|y\|} = \frac{\|X\|\,\|\beta\|}{\|X\beta\|}, \tag{8}$$

and

$$\kappa^{\text{fwd}}(X) = \max_{\beta\in\mathbb{R}^p} \eta^{\text{fwd}}(X,\beta) = \|X\|\left\|X^\dagger\right\| = \kappa(X). \tag{9}$$

Proof It follows from (5) that $\delta y = X\delta\beta$ and thus

$$\|\delta y\| = \|X\delta\beta\| \leq \|X\|\,\|\delta\beta\| = \|X\|\,\|\beta\|\,\Delta\beta,$$

from (7). The result (8) follows by division of this inequality by $\|y\| = \|X\beta\|$. Furthermore,

$$\max_{\delta\beta,\beta\in\mathbb{R}^p} \|\delta y\| = \max_{\beta\in\mathbb{R}^p} \|X\|\,\|\beta\|\,\Delta\beta,$$

and it follows by definition of the forward problem that y lies in the column space of X. The vector β is therefore equal to $X^\dagger y$, which yields

$$\max_{\delta\beta,\beta\in\mathbb{R}^p} \|\delta y\| = \max_{\beta\in\mathbb{R}^p} \|X\|\,\|\beta\|\,\Delta\beta = \|X\|\left\|X^\dagger\right\|\|y\|\,\Delta\beta,$$

and the condition number (9) follows. □

Theorem 3.2 The effective condition number $\eta^{\text{inv}}(X,y)$ of the inverse problem (6) is

$$\eta^{\text{inv}}(X,y) = \max_{\delta y\in\mathbb{R}^n} \frac{\Delta\beta_{\text{LS}}}{\Delta y} = \frac{\left\|X^\dagger\right\|\|y\|}{\|\beta_{\text{LS}}\|} = \frac{\left\|X^\dagger\right\|\|y\|}{\|X^\dagger y\|}. \tag{10}$$

Proof It follows from (6) that $\delta\beta_{\text{LS}} = X^\dagger\delta y$ and thus from (7),

$$\|\delta\beta_{\text{LS}}\| = \left\|X^\dagger\delta y\right\| \leq \left\|X^\dagger\right\|\|\delta y\| = \left\|X^\dagger\right\|\|y\|\,\Delta y.$$

The division of this inequality by $\|\beta_{\text{LS}}\| = \left\|X^{\dagger}y\right\|$ yields the result (10). □

The expressions (8), (9) and (10) assume simpler forms if the 2-norm is used. If the SVD of X is $U\Sigma V^T$ where U and V are orthogonal matrices, and

$$\Sigma = \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} \in \mathbb{R}^{n\times p}, \qquad \Sigma_1 = \text{diag}\begin{bmatrix} \sigma_1 & \sigma_2 & \cdots & \sigma_p \end{bmatrix} \in \mathbb{R}^{p\times p}, \tag{11}$$

then

$$\eta_2^{\text{fwd}}(X,\beta) = \frac{\sigma_1 \|d\|_2}{\|\Sigma d\|_2}, \qquad d = V^T\beta, \tag{12}$$

$$\kappa_2^{\text{fwd}}(X) = \kappa_2(X) = \frac{\sigma_1}{\sigma_p}, \tag{13}$$

$$\eta_2^{\text{inv}}(X,y) = \frac{\|c\|_2}{\sigma_p \|\Sigma^{\dagger}c\|_2}, \qquad c = U^T y. \tag{14}$$

An expression for the condition number $\kappa_2^{\text{inv}}(X)$ of the inverse problem is derived in Theorem 3.3 [13, Theorem 4.2].

Theorem 3.3 Equation (14) can be written as

$$\eta_2^{\text{inv}}(X,y) = \kappa_2(X)\left(\frac{\sum_{i=1}^n c_i^2}{\sum_{i=1}^p \left(\frac{\sigma_1}{\sigma_i}\right)^2 c_i^2}\right)^{\frac{1}{2}}, \tag{15}$$

and thus

$$\eta_2^{\text{inv}}(X,y) \leq \kappa_2(X)\left(\frac{\sum_{i=1}^n c_i^2}{\sum_{i=1}^p c_i^2}\right)^{\frac{1}{2}} = \frac{\kappa_2(X)}{\cos\theta} = \kappa_2^{\text{inv}}(X),$$

where $\kappa_2(X)$ is defined in (13), $\kappa_2^{\text{inv}}(X)$ is the condition number of the inverse problem and

$$\cos\theta = \frac{\|X\beta\|_2}{\|y\|_2} = \frac{\left\|XX^{\dagger}y\right\|_2}{\|y\|_2} = \left(\frac{\sum_{i=1}^p c_i^2}{\sum_{i=1}^n c_i^2}\right)^{\frac{1}{2}}. \tag{16}$$

□

The geometric interpretation of the expression for $\cos\theta$ in (16) requires that the orthogonal projection of a vector onto a subspace be considered. Specifically, the left singular matrix U from the SVD of X is partitioned into two matrices, U_1 and U_2,

$$U = \begin{bmatrix} U_1 \ U_2 \end{bmatrix}, \qquad U_1 \in \mathbb{R}^{n\times p}, \qquad U_2 \in \mathbb{R}^{n\times(n-p)}, \tag{17}$$

where $U_1 U_1^T y$ is the orthogonal projection of y onto $\mathcal{R}(X)$. The expression for c in (14) is written as

$$c = \begin{bmatrix} \bar{c}_1 \\ \bar{c}_2 \end{bmatrix} = \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} y, \tag{18}$$

where $\bar{c}_1 \in \mathbb{R}^p$ and $\bar{c}_2 \in \mathbb{R}^{n-p}$, and thus from (16),

$$\cos^2\theta = \frac{\bar{c}_1^T \bar{c}_1}{c^T c} = \frac{\left\|U_1^T y\right\|_2^2}{\|y\|_2^2} = \frac{y^T(U_1 U_1^T y)}{y^T y}.$$

It follows that as $\cos\theta \to 0$, the angle between y and its orthogonal projection onto $\mathcal{R}(X)$ approaches 90 degrees. This condition should be avoided because it implies, with respect to variable selection, that the chosen predictors are not good parameters for defining the output of the linear system, and with respect to regression, it implies that the chosen basis functions (which are contained in X) are not suitable for representing the data y.

3.2 An uncertainty principle

It is shown in this section that the effective condition numbers of the forward and inverse problems satisfy an uncertainty principle because their product is a function of X, or a function of X and y, depending on the order in which the forward and inverse problems are solved. Theorem 3.4 considers the situation in which the forward problem is solved before the inverse problem, which occurs in variable selection, and Theorem 3.5 considers the situation in which the inverse problem is solved before the forward problem.

Theorem 3.4 Let the vector $\beta = \beta_0$ yield the result y_0 of the forward problem, $y_0 = X\beta_0$. If the inverse problem is then solved with this vector y_0, then the product of the effective condition numbers of the forward and inverse problems is

$$\eta^{\text{fwd}}(X, \beta_0)\eta^{\text{inv}}(X, y_0) = \|X\| \left\|X^\dagger\right\| = \kappa(X). \tag{19}$$

Proof Since the forward problem is solved before the inverse problem, y_0 lies in $\mathcal{R}(X)$ and the residual of the solution $\beta_0 = X^\dagger y_0$ of the inverse problem is equal to zero because

$$X^\dagger y_0 = X^\dagger (X\beta_0) = \beta_0. \tag{20}$$

The effective condition numbers of the forward and inverse problems are, from (8) and (10),

$$\eta^{\text{fwd}}(X, \beta_0) = \frac{\|X\| \, \|\beta_0\|}{\|y_0\|} = \frac{\|X\| \, \|\beta_0\|}{\|X\beta_0\|},$$

and

$$\eta^{\text{inv}}(X, y_0) = \frac{\left\|X^\dagger\right\| \|y_0\|}{\|\beta_0\|} = \frac{\left\|X^\dagger\right\| \|y_0\|}{\|X^\dagger y_0\|},$$

and the result (19) follows from (20). □

Theorem 3.5 considers the situation in which the inverse problem is solved before the forward problem. The result of the theorem is stated in the 2-norm because (14) is required for its proof.

Theorem 3.5 Let the vector $y = y_0$ yield the result $\beta_0 = X^\dagger y_0$ of the inverse problem. If the forward problem is then solved for this vector β_0, that is, $y_1 = X\beta_0$, then the product of the effective condition numbers of the forward and inverse problems is

$$\eta_2^{\text{fwd}}(X, \beta_0)\eta_2^{\text{inv}}(X, y_0) = \kappa_2(X) \left(\frac{\|y_0\|_2}{\left\|U_1 U_1^T y_0\right\|_2} \right) = \frac{\kappa_2(X)}{\cos\theta}, \tag{21}$$

where $\kappa_2(X)$ is defined in (13), $\cos\theta$ is defined in (16), U_1 is defined in (17) and $U_1 U_1^T y_0$ is the orthogonal projection of y_0 onto $\mathcal{R}(X)$. If X is square, that is, $n = p$, the result assumes a simpler form because $\cos\theta = 1$,

$$\eta_2^{\text{fwd}}(X, \beta_0)\eta_2^{\text{inv}}(X, y_0) = \kappa_2(X). \tag{22}$$

Proof It follows from (14) that the effective condition number of the inverse problem is

$$\eta_2^{\text{inv}}(X, y_0) = \frac{\left\| U^T y_0 \right\|_2}{\sigma_p \left\| \Sigma^\dagger U^T y_0 \right\|_2}.$$

The solution of the forward problem is $y_1 = X\beta_0$ where $\beta_0 = X^\dagger y_0$, and the effective condition number of y_1 is, from (12),

$$\eta_2^{\text{fwd}}(X, \beta_0) = \frac{\sigma_1 \left\| V^T \beta_0 \right\|_2}{\left\| \Sigma V^T \beta_0 \right\|_2}.$$

The product of these effective condition numbers is

$$\begin{aligned}
\eta_2^{\text{fwd}}(X, \beta_0)\eta_2^{\text{inv}}(X, y_0) &= \frac{\sigma_1}{\sigma_p} \left(\frac{\left\| U^T y_0 \right\|_2}{\left\| \Sigma V^T \beta_0 \right\|_2} \right) \left(\frac{\left\| V^T \beta_0 \right\|_2}{\left\| \Sigma^\dagger U^T y_0 \right\|_2} \right) \\
&= \kappa_2(X) \left(\frac{\left\| y_0 \right\|_2}{\left\| X X^\dagger y_0 \right\|_2} \right) \left(\frac{\left\| V^T \beta_0 \right\|_2}{\left\| \Sigma^\dagger U^T y_0 \right\|_2} \right) \\
&= \kappa_2(X) \left(\frac{\left\| y_0 \right\|_2}{\left\| X X^\dagger y_0 \right\|_2} \right),
\end{aligned}$$

and the results (21) and (22) follow. □

It follows from Theorem 3.4 that if the forward problem is solved before the inverse problem, then the term on the right hand side of the uncertainty principle is equal to $\kappa(X)$. Theorem 3.5 shows, however, that if the inverse problem is solved before the forward problem, then the term on the right hand side of the uncertainty principle may be much larger than $\kappa_2(X)$. It was noted in Section 3.1 that $\cos\theta \approx 1$ in a well formulated problem, and the uncertainty principles (19), (21) and (22) can therefore be combined,

$$\eta_2^{\text{fwd}}(X, \beta_0)\eta_2^{\text{inv}}(X, y_0) \approx \kappa_2(X).$$

This form of the uncertainty principle is independent of the order in which the forward and inverse problems are solved, and it has two important consequences:

(1) If X is well conditioned, then the forward and inverse problems are also well conditioned.
(2) If X is ill conditioned, then three situations arise:
   (a) The forward problem is well conditioned and the inverse problem is ill conditioned.
   (b) The forward problem is ill conditioned and the inverse problem is well conditioned.

(c) The forward and inverse problems satisfy

$$\eta_2^{\text{fwd}}(X,\beta_0) \approx \eta_2^{\text{inv}}(X,y_0) \approx \sqrt{\kappa_2(X)}.$$

An example of the application of the uncertainty principle to regression is in [13, §3].

4 The discrete Picard condition and Tikhonov regularisation

The effective condition numbers of the forward and inverse problems were considered in Section 3, and it is instructive to consider the conditions that are satisfied for the effective condition number of the inverse problem to attain its maximum value because it leads to the discrete Picard condition, the satisfaction of which is required for the correct application of Tikhonov regularisation [6, §4.5]. This application is discussed in Section 4.1, and it is shown in Section 4.2 that this condition also arises in the consideration of the stability of the forward problem in variable selection.

4.1 The discrete Picard condition and the inverse problem

It follows from (14) that $\eta_2^{\text{inv}}(X,y)$ attains its maximum value when

$$\frac{|c_i|}{\sigma_i} \to 0 \qquad \text{as} \qquad i \to p, \tag{23}$$

where $c = \{c_i\}_{i=1}^n$. The singular values σ_i are arranged in non-increasing order and it therefore follows from (23) that $\eta_2^{\text{inv}}(X,y)$ attains its maximum value when the constants $|c_i|$ decay to zero faster than the singular values decay to zero. If the decay (23) is such that $|c_i|/\sigma_i \approx 0, i = k+1,\dots,p$, then $|c_i| \approx 0, i = k+1,\dots,p$, and it follows from (15) that if $\cos\theta \approx 1$, then

$$\begin{aligned}
\eta_2^{\text{inv}}(X,y) &\approx \kappa_2(X) \left(\frac{\sum_{i=1}^k c_i^2 + \sum_{i=p+1}^n c_i^2}{\sigma_1^2 \sum_{i=1}^k \left(\frac{c_i}{\sigma_i}\right)^2} \right)^{\frac{1}{2}} \\
&\approx \kappa_2(X) \left(\frac{\sum_{i=1}^k c_i^2}{\sigma_1^2 \sum_{i=1}^k \left(\frac{c_i}{\sigma_i}\right)^2} \right)^{\frac{1}{2}} \\
&\approx \kappa_2(X).
\end{aligned}$$

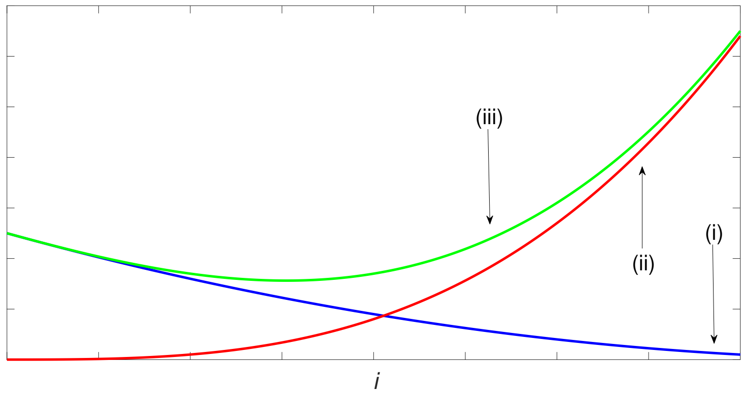

Fig. 1. The ratios (i) $|c_i|/\sigma_i$ —, (ii) $|\delta c_i|/\sigma_i \approx \epsilon/\sigma_i$ — and (iii) $|c_i+\delta c_i|/\sigma_i$ — if the discrete Picard condition is satisfied.

Equation (23) is the discrete Picard condition, which is defined because of its importance.

Definition 4.1 (The discrete Picard condition) The discrete Picard condition is satisfied if the constants $|c_i|, i = 1, \ldots, p$, decay to zero faster than the singular values σ_i decay to zero.

It follows from (14) that $y = Uc$ and thus the dominant components of y lie in the space spanned by the first few columns of U if the discrete Picard condition is satisfied.

It is shown in [13, §4] that $|c_i|/\sigma_i$ is very sensitive to a perturbation $\delta c = U^T \delta y$ because even if the theoretically exact solution satisfies the discrete Picard condition, the perturbed solution does not satisfy this condition,

$$\frac{|c_i + \delta c_i|}{\sigma_i} \nrightarrow 0 \qquad \text{as} \qquad i \to p,$$

where the perturbations δc_i satisfy $|\delta c_i| \approx \epsilon$, $i = 1, \ldots, p$, such that

$$\begin{aligned} |\delta c_i| &\approx \epsilon \ll |c_i|, \quad && i = 1, \ldots, t, \\ |\delta c_i| &\approx \epsilon \gg |c_i|, \quad && i = t+1, \ldots, p. \end{aligned} \tag{24}$$

This error model arises because the constants $|c_i|$ decay to zero faster than the singular values decay to zero and it is assumed that the magnitude $|\delta c_i|$ of the perturbations is approximately constant. Figure 1 shows the variation of the ratios $|c_i|/\sigma_i$, $|\delta c_i|/\sigma_i \approx \epsilon/\sigma_i$ and $|c_i+\delta c_i|/\sigma_i$ with i if the discrete Picard condition is satisfied, and it is seen that the magnitude of the solution of the inverse problem in the presence of noise is governed by the noise ϵ,

$$\|\beta+\delta\beta\|_2^2=\sum_{i=1}^{p}\left(\frac{c_i+\delta c_i}{\sigma_i}\right)^2\approx\sum_{i=t+1}^{p}\left(\frac{\delta c_i}{\sigma_i}\right)^2\approx\sum_{i=t+1}^{p}\left(\frac{\epsilon}{\sigma_i}\right)^2\approx\left(\frac{\epsilon}{\sigma_p}\right)^2, \tag{25}$$

which confirms that the solution of the inverse problem is ill conditioned. It follows that the satisfaction of the discrete Picard condition by the exact solution does not imply that the solution obtained with perturbed data satisfies this condition. Furthermore,

$$\eta_2^{\text{inv}}(X,y+\delta y)=\frac{\|c+\delta c\|_2}{\sigma_p\left\|\Sigma^{\dagger}(c+\delta c)\right\|_2}\approx\frac{\|c\|_2}{\sigma_p\left(\frac{|c_p+\delta c_p|}{\sigma_p}\right)}\approx\frac{\|c\|_2}{\sigma_p\left(\frac{\epsilon}{\sigma_p}\right)}=\frac{\|c\|_2}{\epsilon}, \tag{26}$$

which shows that the effective condition number of the inverse problem is ill conditioned if the discrete Picard condition is satisfied.

Equations (25) and (26) show that if the exact solution of the inverse problem satisfies the discrete Picard condition, then the solution and its effective condition number in the presence of noise are dominated by noise. It is shown, however, in Section 4.3 that the satisfaction of the discrete Picard condition guarantees that Tikhonov regularisation yields an approximate solution of an ill conditioned LS problem that satisfies the conditions stated in Section 1 (the approximate solution is numerically stable and the regularisation error is small). The correct application of Tikhonov regularisation is therefore based on prior knowledge of properties of the theoretically exact solution. For example, an image is dominated by spectral components of low frequency, which implies that the discrete Picard condition is satisfied, and thus a blurred image can be restored to a deblurred form by the application of Tikhonov regularisation [8, pp. 67-69].

The importance of the discrete Picard condition can also be seen by considering other forms of the ratio ${}^{|c_i|}/_{\sigma_i}$. Specifically, if

$$|c_i|\approx\sigma_i,\qquad i=1,\ldots,p, \tag{27}$$

then ${}^{|c_i|}/_{\sigma_i}\approx 1$, and furthermore, the perturbation model (24) is appropriate because the constants $|c_i|$ decay to zero as i increases. It is easily established that $\eta_2^{\text{inv}}(X,y)$ cannot be computed reliably for the model (27) because it is sensitive to the perturbations δc_i.

If the constants $|c_i|$ satisfy

$$|c_{i+1}|\gg|c_i|,\qquad i=1,\ldots,p-1, \tag{28}$$

then $|c_i|/\sigma_i$ increases monotonically as i increases, and if the perturbations δc_i satisfy (24), then $\eta_2^{\text{inv}}(X, y)$ can be computed reliably because the effect of the perturbations is small [13, §4].

The models (23), (27) and (28) are useful for analysing the solution of the inverse problem, and it is shown in Section 4.3 that Tikhonov regularisation cannot be used if the data satisfy (27) or (28) because these models lead to large regularisation errors.

4.2 The discrete Picard condition and the forward problem

It was shown in Section 4.1 that the satisfaction of the discrete Picard condition implies that the solution and effective condition number of the inverse problem are ill conditioned. This analysis is extended in this section to the forward problem, whose output is equal to the input of the inverse problem in variable selection, and it is shown that $\eta_2^{\text{fwd}}(X, \beta)$ can be written in a form that includes the term $|c_i|/\sigma_i$ that defines the discrete Picard condition. It therefore follows that the stability and instability issues that were considered in Section 4.1 for the inverse problem must also be considered for the forward problem.

It follows from (11), (12), (14) and (18) that

$$c = \begin{bmatrix} \bar{c}_1 \\ \bar{c}_2 \end{bmatrix} = U^T y = U^T X\beta = U^T(U\Sigma V^T)(Vd) = \Sigma d = \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} d,$$

and thus

$$d = \Sigma_1^{-1}\bar{c}_1 = \begin{bmatrix} \frac{c_1}{\sigma_1} & \frac{c_2}{\sigma_2} & \cdots & \frac{c_p}{\sigma_p} \end{bmatrix}^T. \tag{29}$$

This equation unites the forward and inverse problems because (12) and (14) show that d and c arise in the effective condition numbers of the forward and inverse problems, respectively. Also, the equation allows the expression for the effective condition number of the forward problem (12) to be written as

$$\eta_2^{\text{fwd}}(X, \beta) = \frac{\sigma_1 \left\| d \right\|_2}{\left\| \Sigma d \right\|_2} = \frac{\sigma_1 \left\| \Sigma_1^{-1}\bar{c}_1 \right\|_2}{\left\| \Sigma\Sigma_1^{-1}\bar{c}_1 \right\|_2} = \sigma_1 \left(\frac{\sum_{i=1}^{p} \left(\frac{c_i}{\sigma_i}\right)^2}{\sum_{i=1}^{p} c_i^2} \right)^{\frac{1}{2}}, \tag{30}$$

and thus the term c_i/σ_i that defines the solution and effective condition number of the inverse problem also defines the effective condition number of the forward problem. It is therefore necessary to consider the stability of (30) for the conditions $\eta_2^{\text{fwd}}(X,\beta) \approx 1$ and $\eta_2^{\text{fwd}}(X,\beta) \approx \kappa_2(X)$. These situations are considered in Theorems 4.1 and 4.2 respectively.

Theorem 4.1 If the forward problem (5) is well conditioned, and the inverse problem (6) is ill conditioned and $\cos\theta \approx 1$, then $\eta_2^{\text{fwd}}(X,\beta) \approx 1$ and $\eta_2^{\text{inv}}(X,y) \approx \kappa_2(X) \gg 1$. Furthermore, the effective condition numbers of the forward and inverse problems are ill conditioned.

Proof The condition $\eta_2^{\text{inv}}(X,y) \approx \kappa_2(X)$ implies that the discrete Picard condition is satisfied and $\cos\theta \approx 1$. If β is perturbed to $\beta + \delta\beta$, then it follows from (30) that

$$\eta_2^{\text{fwd}}(X,\beta+\delta\beta) = \sigma_1 \left(\frac{\sum_{i=1}^{p} \left(\frac{c_i+\delta c_i}{\sigma_i}\right)^2}{\sum_{i=1}^{p}(c_i+\delta c_i)^2} \right)^{\frac{1}{2}} \approx \sigma_1 \left(\frac{|c_p + \delta c_p|}{\sigma_p} \right) \frac{1}{\|c\|_2},$$

and thus $\eta_2^{\text{fwd}}(X,\beta)$ is ill conditioned because

$$\eta_2^{\text{fwd}}(X,\beta+\delta\beta) \approx \left(\frac{\epsilon}{\|c\|_2} \right) \kappa_2(X).$$

The ill conditioned nature of $\eta_2^{\text{inv}}(X,y)$ when it is approximately equal to $\kappa_2(X)$ is established in (26), and in accordance with the uncertainty principle, the product of these effective condition numbers is approximately equal to $\kappa_2(X)$. It is noted that the forward problem is well conditioned and its effective condition number is ill conditioned, but the inverse problem and its effective condition number are ill conditioned. □

Theorem 4.2 If the forward problem (5) is ill conditioned and its effective condition number satisfies $\eta_2^{\text{fwd}}(X,\beta) \approx \kappa_2(X)$, then the inverse problem (6) is well conditioned and the effective condition numbers of the forward and inverse problems are well conditioned.

Proof It follows from the uncertainty principle that the inverse problem is well conditioned and thus (28) is satisfied,

$$\eta_2^{\text{inv}}(X,y) = \frac{\|c\|_2}{\sigma_p \|\Sigma^\dagger c\|_2} \approx \frac{|c_p|}{\sigma_p \left(\frac{|c_p|}{\sigma_p}\right)} = 1.$$

Also, it follows from (28) and (29) that $\|d\|_2 \approx {}^{|c_p|}/_{\sigma_p}$ and thus the forward problem is ill conditioned,

$$\eta_2^{\text{fwd}}(X,\beta) = \sigma_1 \frac{\|d\|_2}{\|\Sigma d\|_2} \approx \sigma_1 \left(\frac{|c_p|}{\sigma_p}\right)\frac{1}{|c_p|} = \kappa_2(X).$$

If perturbed data $c_i+\delta c_i$ are considered and $\cos\theta \approx 1$ for this perturbed data, where the perturbations δc_i satisfy (24), then

$$\eta_2^{\text{inv}}(X, y+\delta y) \approx \frac{1}{\sigma_p}\left(\frac{\sum_{i=1}^{p}(c_i+\delta c_i)^2}{\sum_{i=1}^{p}\left(\frac{c_i+\delta c_i}{\sigma_i}\right)^2}\right)^{\frac{1}{2}} \approx \frac{1}{\sigma_p}\left(\frac{|c_p|}{\left(\frac{|c_p|}{\sigma_p}\right)}\right) = 1,$$

and

$$\eta_2^{\text{fwd}}(X,\beta+\delta\beta) = \frac{\sigma_1\|d+\delta d\|_2}{\|\Sigma(d+\delta d)\|_2} = \frac{\sigma_1\left\|\Sigma_1^{-1}(\bar{c}_1+\delta\bar{c}_1)\right\|_2}{\left\|\Sigma\Sigma_1^{-1}(\bar{c}_1+\delta\bar{c}_1)\right\|_2} \approx \sigma_1\left(\frac{|c_p|}{\sigma_p}\right)\frac{1}{|c_p|},$$

and thus

$$\eta_2^{\text{fwd}}(X,\beta+\delta\beta) \approx \kappa_2(X).$$

It follows that the effective condition numbers of the forward and inverse problems are well conditioned. □

Theorems 4.1 and 4.2 are summarised in Table 1 and it is seen that a distinction is made between the stability/instability of a problem, and the stability/instability of its effective condition number.

	Fwd. problem well cond.	Fwd. problem ill cond.
Inverse problem well cond.	−−−	$\eta_2^{\text{fwd}}(X,\beta)$ and $\eta_2^{\text{inv}}(X,y)$ are well conditioned
Inverse problem ill cond.	$\eta_2^{\text{fwd}}(X,\beta)$ and $\eta_2^{\text{inv}}(X,y)$ are ill conditioned	−−−

Table 1
The stability of the forward and inverse problems, and the stability of their effective condition numbers, if $\kappa_2(X) \gg 1$.

4.3 Tikhonov regularisation

This section considers Tikhonov regularisation, which requires that a constraint be imposed on $\|\beta\|_2^2$ and leads to the equation

$$(X^TX + \lambda I)\beta(\lambda) = X^Ty, \qquad \lambda \geq 0, \tag{31}$$

where λ is the regularisation parameter. This equation is derived from (4) by setting $\lambda_1 = 0$ and $\lambda_2 = \lambda$. Methods for the computation of the optimal value of λ are discussed in Section 4.4, but it is considered a known constant in this section. The solution of (31) is

$$\beta(\lambda) = V(\Sigma^T\Sigma + \lambda I)^{-1}\Sigma^T c = V\left((\Sigma^T\Sigma + \lambda I)^{-1}\Sigma^T\Sigma\right)\Sigma^\dagger c = VF(\lambda)\Sigma^\dagger c, \tag{32}$$

where $F(\lambda)$ is a square diagonal matrix of order p whose non-zero entries $f_i(\lambda), i = 1, \ldots, p$, are the filter factors of X,

$$F(\lambda) = (\Sigma^T\Sigma + \lambda I)^{-1}\Sigma^T\Sigma = \text{diag}\ \{f_i(\lambda)\}_{i=1}^p = \text{diag}\left\{\frac{\sigma_i^2}{\sigma_i^2 + \lambda}\right\}_{i=1}^p,$$

and $\beta(0) = V\Sigma^\dagger c$ is equal to β_{LS}, which is defined in (3). The effect of λ can be quantified by assuming there exists an index t, $t < p$, such that $\lambda \approx \sigma_t^2$, in which case the filter factors $f_i(\lambda) \approx f_i(\sigma_t^2)$ satisfy

$$f_i(\lambda) = \begin{cases} \frac{\sigma_i^2}{\sigma_i^2+\lambda} \approx 1, & i < t, \\ \frac{\sigma_i^2}{\sigma_i^2+\lambda} \approx \frac{1}{2}, & i = t, \\ \frac{\sigma_i^2}{\sigma_i^2+\lambda} \approx 0, & i > t, \end{cases} \qquad \lambda \approx \sigma_t^2. \tag{33}$$

It therefore follows from (32) that the solution of (31) can be written as

$$\beta(\lambda) = \sum_{i=1}^p f_i(\lambda)\left(\frac{c_i}{\sigma_i}\right)v_i \approx \sum_{i=1}^t f_i(\lambda)\left(\frac{c_i}{\sigma_i}\right)v_i, \qquad \lambda \approx \sigma_t^2, \tag{34}$$

where v_i is the ith column of V. The filter factors of the 8×8 Hilbert matrix are considered in Example 4.1.

Example 4.1 Figure 2 shows the filter factors $f_i(\lambda)$ of the 8×8 Hilbert matrix H for four values of λ. It is seen that they decay rapidly to zero, and thus the effect of the small singular values of H on $\beta(\lambda)$ is also reduced to zero because of this decay of the filter factors. □

The filter factors must decay to zero for the suppression of the small singular values of X from $\beta(\lambda)$, which is important for Tikhonov regularisation. Sections 5 and 6 consider the situation in which X is a design matrix with

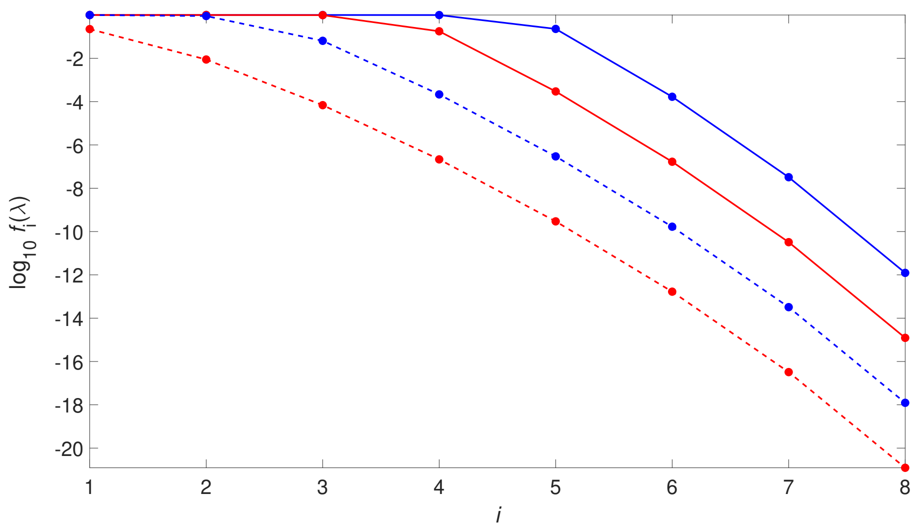

Fig. 2. Filter factors of the 8×8 Hilbert matrix for $\lambda = 10^{-8}$ —, $\lambda = 10^{-5}$ —, $\lambda = 10^{-2}$ – – and $\lambda = 10$ – –, for Example 4.1.

correlated covariates and it will be shown that this decay of the filter factors is not satisfied, which has consequences for the effectiveness of Tikhonov regularisation for variable selection.

The regularisation error, that is, the error between $\beta(\lambda)$ and $\beta(0) = \beta_{\text{LS}} = X^\dagger y = V\Sigma^\dagger c$, as a function of λ is considered in Theorem 4.3 [13, §5].

Theorem 4.3 The regularisation error $\Delta(\lambda)$ between $\beta(\lambda)$ and $\beta(0)$ is

$$\Delta(\lambda) = \frac{\|\beta(\lambda) - \beta(0)\|_2}{\|\beta(0)\|_2} = \lambda \left(\frac{\sum_{i=1}^p \left(\frac{c_i}{\sigma_i}\right)^2 \frac{1}{(\sigma_i^2+\lambda)^2}}{\sum_{i=1}^p \left(\frac{c_i}{\sigma_i}\right)^2} \right)^{\frac{1}{2}}.$$

Let the filter factors satisfy (33). Then the regularisation error $\Delta(\lambda \approx \sigma_t^2)$, is a function of the ratio $|c_i|/\sigma_i$:

(i) If the discrete Picard condition (23) is satisfied, then

$$\Delta(\lambda \approx \sigma_t^2) \approx \left(\frac{\sigma_t}{\sigma_1}\right)^2 \ll 1. \tag{35}$$

(ii) If (27) is satisfied, then

$$\Delta(\lambda \approx \sigma_t^2) \approx \left(\frac{p-t}{p}\right)^{\frac{1}{2}} < 1.$$

(iii) If (28) is satisfied, then

$$\Delta(\lambda \approx \sigma_t^2) \approx \frac{\sigma_t^2}{\sigma_p^2 + \sigma_t^2} \approx 1.$$

□

Theorem 4.3 shows that the regularisation error is large if $c = U^T y$ satisfies (27) or (28), and thus regularisation must not be applied if these conditions are satisfied. Regularisation is, however, applied when the discrete Picard condition is satisfied, and Theorem 4.3 shows that the regularisation error (35) is small in this circumstance.

It was stated in Section 1 that a regularised solution must be numerically stable and have a small regularisation error. The regularisation error was considered in Theorem 4.3, and it is shown in Theorem 4.4 that the satisfaction of the discrete Picard condition guarantees that the regularised solution is numerically stable [13, §5].

Theorem 4.4 Let the relative errors $\Delta\beta(\lambda)$ and Δy be defined in the 2-norm. The effective condition number of $\beta(\lambda)$ is

$$\eta_2^{\text{inv}}(X, y, \lambda) = \max_{\delta y \in \mathbb{R}^n} \frac{\Delta\beta(\lambda)}{\Delta y} = \frac{\left\| (\Sigma^T\Sigma + \lambda I)^{-1}\Sigma^T \right\|_2 \|c\|_2}{\left\| (\Sigma^T\Sigma + \lambda I)^{-1}\Sigma^T c \right\|_2},$$

and if the discrete Picard condition is satisfied, the filter factors satisfy (33) and $\cos\theta \approx 1$, then

$$\eta_2^{\text{inv}}(X, y, \lambda = \sigma_t^2) \approx \gamma \left(\frac{\sigma_1}{\sigma_t}\right) < \eta_2^{\text{inv}}(X, y, \lambda = 0) \approx \frac{\sigma_1}{\sigma_p}, \quad \frac{1}{2} \leq \gamma \leq 1. \quad (36)$$

□

Equation (36) is expected because it follows from (33) that the filter factors retain the first t singular values $\sigma_i, i = 1, \ldots, t$, of X, and the last $p-t$ singular values $\sigma_i, i = t+1, \ldots, p$, of X are removed from the solution.

4.4 Methods for computing the optimal value of λ

Two popular methods for the determination of the optimal value λ_{opt} of λ are the GCV [1], [3], [4, Chapter 15] and [6, §7.4], and the L-curve [4, Chapter 15], [5] and [6, §4.5], and both methods require that the discrete Picard condition be satisfied [6, §7.4] and [6, §7.5.1]. Furthermore, the computation of λ_{opt} using the GCV requires that the noise be white [11] and it is necessary to determine the value of λ for which a function attains its minimum value, but this minimum is very often shallow, which makes its accurate computation difficult. Computational problems with the L-curve are discussed in [6, §7.5.2].

5 Variable selection with correlated covariates

The forward and inverse problems discussed in this paper occur in variable selection, and the situation in which X is a design matrix with correlated covariates is considered in this section and Section 6. The number of rows n of X is equal to the number of observations, and the number of columns p of X is equal to $d+1$, where d is the number of predictors. The columns of X are $\mathbf{1}$, the column vector all of whose entries are one, and the covariates $\mathbf{x}_i$, $i = 1, \ldots, d$,

$$X = \begin{bmatrix} \mathbf{1} & \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_d \end{bmatrix} \in \mathbb{R}^{n \times p}, \qquad \mathbf{1}, \mathbf{x}_i \in \mathbb{R}^n. \tag{37}$$

The covariates are jointly Gaussian and marginally distributed as $\mathcal{N}(0, \sigma^2)$ with correlation coefficient r,

$$r = \operatorname{cor}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\mathbf{x}_i^T \mathbf{x}_j}{\|\mathbf{x}_i\|_2 \|\mathbf{x}_j\|_2}, \qquad i, j = 1, \ldots, d.$$

It follows from (37) that

$$X^T X = \begin{bmatrix} n & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{x}_1^T \mathbf{x}_1 & \mathbf{x}_1^T \mathbf{x}_2 & \cdots & \mathbf{x}_1^T \mathbf{x}_d \\ 0 & \mathbf{x}_2^T \mathbf{x}_1 & \mathbf{x}_2^T \mathbf{x}_2 & \cdots & \mathbf{x}_2^T \mathbf{x}_d \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \mathbf{x}_d^T \mathbf{x}_1 & \mathbf{x}_d^T \mathbf{x}_2 & \cdots & \mathbf{x}_d^T \mathbf{x}_d \end{bmatrix},$$

where the variance σ^2 of the jth column, $j = 2, \ldots, p$, of X is

$$\sigma^2 = \frac{E\left\{\mathbf{x}_i^T \mathbf{x}_i\right\}}{n-1}, \qquad i = 1, \ldots, d,$$

and thus

$$\|\mathbf{x}_i\|_2 = \sigma\sqrt{n-1}, \qquad i = 1, \ldots, d.$$

It follows that

$$\begin{aligned}
\tilde{X}^T\tilde{X} &= E\left\{X^T X\right\} \\
&= \begin{bmatrix}
n & 0 & 0 & 0 & \cdots & 0 \\
0 & (n-1)\sigma^2 & r(n-1)\sigma^2 & r(n-1)\sigma^2 & \cdots & r(n-1)\sigma^2 \\
0 & r(n-1)\sigma^2 & (n-1)\sigma^2 & r(n-1)\sigma^2 & \cdots & r(n-1)\sigma^2 \\
0 & r(n-1)\sigma^2 & r(n-1)\sigma^2 & (n-1)\sigma^2 & \cdots & r(n-1)\sigma^2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & r(n-1)\sigma^2 & r(n-1)\sigma^2 & r(n-1)\sigma^2 & \cdots & (n-1)\sigma^2
\end{bmatrix} \\
&= \begin{bmatrix}
n & 0 & \cdots & 0 \\
0 & & & \\
\vdots & & \sigma^2(n-1)\left[(1-r)I + rJ\right] & \\
0 & & &
\end{bmatrix},
\end{aligned} \tag{38}$$

where all the entries of $J \in \mathbb{R}^{d\times d}$ are one, and $(1-r)I + rJ$ is a symmetric matrix whose diagonal entries are one and all other entries are equal to r. This expression for $\tilde{X}^T\tilde{X}$ enables its singular values to be calculated [9]. In particular, the eigenvalues of rJ in (38) are

$$\mu_1 = rd, \qquad \mu_2 = \mu_3 = \cdots = \mu_d = 0,$$

and thus the eigenvalues of $(1-r)I + rJ$ are

$$\mu_1 = rd - r + 1, \qquad \mu_2 = \mu_3 = \cdots = \mu_d = 1 - r.$$

The singular values of $\tilde{X}$ are therefore

$$\sigma_1 = \sqrt{n}, \qquad \sigma_2 = \sigma\sqrt{(n-1)(rd-r+1)}, \qquad \sigma_3 = \sigma\sqrt{(n-1)(1-r)}, \tag{39}$$

where $\sigma_2 > \sigma_3$ and the multiplicity of σ_3 is $p-2$.

6 Examples

This section considers two examples of regularisation and one example of variable selection. Examples 6.1 and 6.2 consider regularisation, and in particular, the forward problem is well conditioned and the inverse problem is ill conditioned in Example 6.1, and the forward problem is ill conditioned and the inverse problem is well conditioned in Example 6.2. These problems are

$$\begin{aligned}&\text{Forward problem:} \quad y_0 = X\beta_0 + \epsilon, \\ &\text{Inverse problem:} \quad \beta(0) = \beta_{\text{LS}} = X^\dagger y_0 = X^\dagger(X\beta_0 + \epsilon) = \beta_0 + X^\dagger \epsilon, \end{aligned} \quad (40)$$

where ϵ is a vector of uniformly distributed random variables, $\beta(0) = \beta_{\text{LS}}$ is the solution of the inverse problem for $\lambda = 0$, and the signal-to-noise ratio (SNR) of the forward problem is $\|X\beta_0\|_2 / \|\epsilon\|_2$.

Example 6.3 considers variable selection, for which Algorithm 1 with $N = 1$ is implemented. The vector β_0 is dense in Examples 6.1 and 6.2, but it is sparse in Example 6.3.

Example 6.1 Let the design matrix X be of order 150×11, the correlation coefficient r be equal to 0.9999 and the variance σ^2 be equal to 1. The condition number of X is $\kappa_1(X) = 1269$ and $\kappa_2(X) = 316.2$, and the vector β_0 is such that the effective condition numbers in the absence of noise ($\epsilon = 0$ in (40)) are

$$\begin{aligned} \eta_1^{\text{fwd}}(X, \beta_0) &= 1.2460, \qquad & \eta_1^{\text{inv}}(X, y_0) &= 1018.1, \\ \eta_2^{\text{fwd}}(X, \beta_0) &= 1.0003, \qquad & \eta_2^{\text{inv}}(X, y_0) &= 316.11, \end{aligned}$$

where β_0 and y_0 are defined in (40). It follows that the forward problem is well conditioned and the inverse problem is ill conditioned.

Noise ϵ was added to $X\beta_0$, as shown in (40), and scaled such that the SNR is equal to 10. Figure 3 shows the variation of $\log_{10} |d_i| \, \sigma_i$, $\log_{10} |c_i|/\sigma_i$, and $\log_{10} \sigma_i$ with i, where $d = \{d_i\}$ and $c = \{c_i\}$ are defined in (12) and (14) respectively, and σ_i are the singular values of X. It is seen that X has two singular values of unit multiplicity and one singular value of multiplicity $p - 2 = 9$, which follows from (39). The figure also shows that the dominant components of $|d_i| \, \sigma_i$ and $|c_i|/\sigma_i$ are defined by $i = 1$ and $i = 2$, and that the other components of these functions are much smaller, by a few orders of magnitude. Tikhonov regularisation was used to regularise the inverse problem, and the optimal value of λ was determined by the GCV, which requires the evaluation of a function $G(\lambda)$, and the L-curve. The MATLAB package Regularization Tools [7] was used for these computations, and the L-curve and GCV are shown in

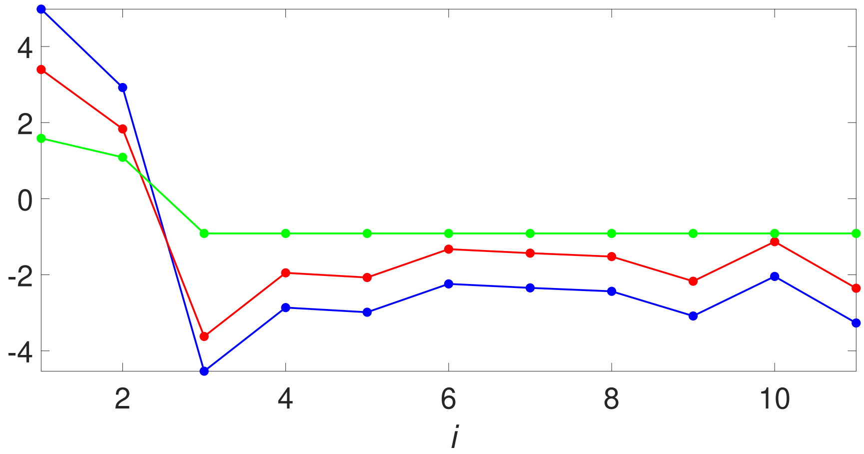

Fig. 3. The variation of $\log_{10} |d_i|\, \sigma_i$ $\bullet$, $\log_{10} {|c_i|}/{\sigma_i}$ $\bullet$ and $\log_{10} \sigma_i$ $\bullet$ with i, for Example 6.1.

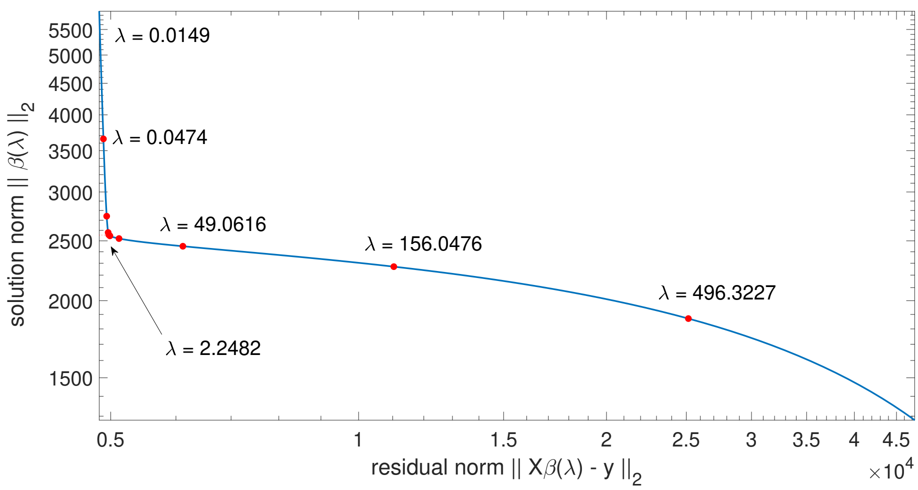

Fig. 4. The L-curve and the optimal value of the regularisation parameter, $\lambda_{\text{opt}}^{\text{Lcurve}} = 2.2482$, for Example 6.1.

Figures 4 and 5 respectively. The optimal value of the regularisation parameter from the L-curve, $\lambda_{\text{opt}}^{\text{Lcurve}}$, is the value of λ in the corner, which is the point on the curve at which the curvature is a maximum. The optimal value of the regularisation parameter from the GCV, $\lambda_{\text{opt}}^{\text{GCV}}$, is the value of λ for which $G(\lambda)$ attains its minimum value. Figure 5 highlights a problem with the GCV, specifically, the function $G(\lambda)$ is almost flat in the neighbourhood of $\lambda_{\text{opt}}^{\text{GCV}}$, which makes its computation difficult. It is noted that the optimal values of λ obtained from these methods differ by about three orders of magnitude.

The correct application of Tikhonov regularisation requires that the discrete Picard condition be satisfied, such that the filter factors $f_i(\lambda)$ remove the terms in the solution $\beta(0)$ of the inverse problem that are defined by the small

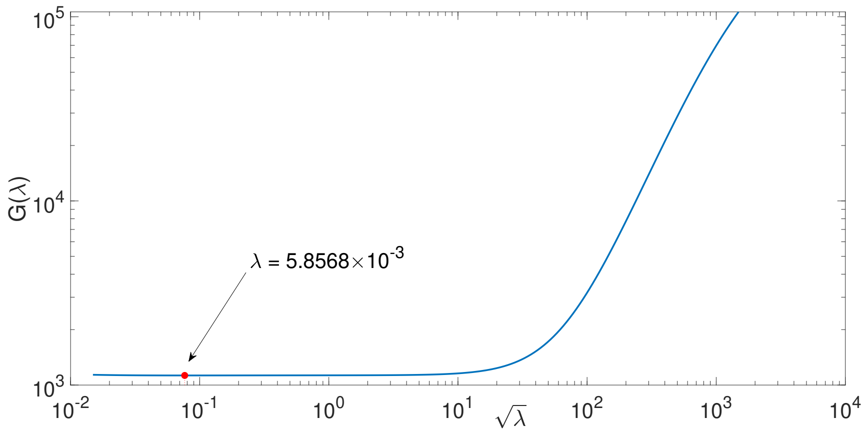

Fig. 5. The GCV function $G(\lambda)$ against $\sqrt{\lambda}$ and the optimal value of the regularisation parameter, $\lambda_{\text{opt}}^{\text{GCV}} = 5.8568 \times 10^{-3}$, for Example 6.1.

singular values, that is, the terms that corrupt $\beta(0)$, as shown in (34). This removal requires that the filter factors decay to zero, but the error in the regularised solution may be large if this decay condition is not satisfied. The filter factors for the optimal regularisation parameters for the L-curve and GCV, $\lambda_{\text{opt}}^{\text{Lcurve}} = 2.2482$ and $\lambda_{\text{opt}}^{\text{GCV}} = 5.8568 \times 10^{-3}$ respectively, are shown in Figure 6 and it is seen that their values computed from $\lambda_{\text{opt}}^{\text{GCV}}$ are

$$\begin{aligned} f_i(\lambda_{\text{opt}}^{\text{GCV}}) &\approx 1, & i &= 1, 2, \\ f_i(\lambda_{\text{opt}}^{\text{GCV}}) &\approx 10^{-0.144} = 0.718, & i &= 3, \ldots, 11, \end{aligned}$$

and the ratio 0.718 of the minimum value of the filter factors to the maximum value of the filter factors shows that these filter factors are not effective in removing the small singular values from $\beta(0)$. The filter factors computed from $\lambda_{\text{opt}}^{\text{Lcurve}}$ are more effective in the removal of these singular values because

$$\begin{aligned} f_i(\lambda_{\text{opt}}^{\text{Lcurve}}) &\approx 1, & i &= 1, 2, \\ f_i(\lambda_{\text{opt}}^{\text{Lcurve}}) &\approx 10^{-2.182} = 6.58 \times 10^{-3}, & i &= 3, \ldots, 11, \end{aligned}$$

and thus the ratio of the minimum value of the filter factors to the maximum value of the filter factors is equal to 6.58×10^{-3}, which is much smaller than its value of 0.718 for $\lambda_{\text{opt}}^{\text{GCV}}$. The filter factors do not decay to zero and they therefore differ from the filter factors shown in Figure 2 for the 8×8 Hilbert matrix. This property of the filter factors of X follows because its singular values do not decay to zero but level off at a constant value σ_i, $i = 3, \ldots, 11$, as shown in Figure 3.

This error, and the error e^{LS} of the LS solution $\beta_{\text{LS}} = \beta(0)$, are

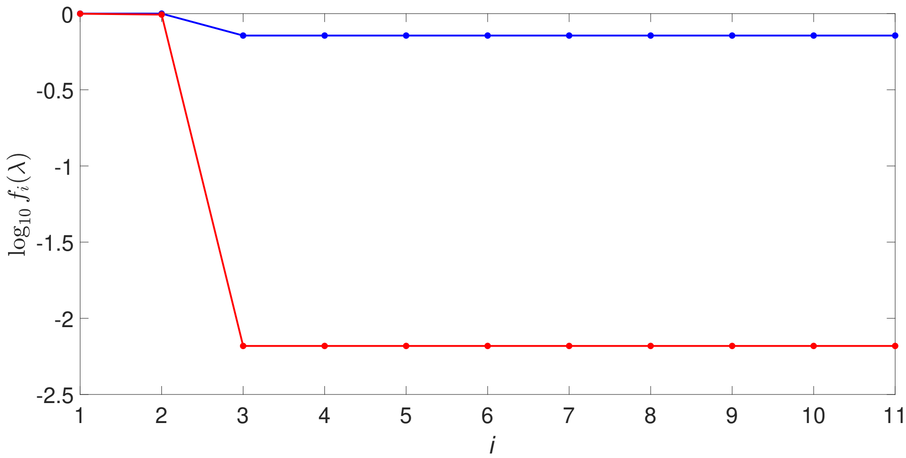

Fig. 6. The filter factors for the optimal regularisation parameter for the L-curve, $\lambda = \lambda_{\text{opt}}^{\text{Lcurve}} = 2.2482$ $\bullet$, and the GCV, $\lambda = \lambda_{\text{opt}}^{\text{GCV}} = 5.8568 \times 10^{-3}$ $\bullet$, for Example 6.1.

$$
\begin{aligned}
e^{\text{LS}} &= \frac{\|\beta_{\text{LS}} - \beta_0\|_2}{\|\beta_0\|_2} = 4.3947, \\
e^{\text{Lcurve}} &= \frac{\left\|\beta(\lambda_{\text{opt}}^{\text{Lcurve}}) - \beta_0\right\|_2}{\|\beta_0\|_2} = 0.2675, \\
e^{\text{GCV}} &= \frac{\left\|\beta(\lambda_{\text{opt}}^{\text{GCV}}) - \beta_0\right\|_2}{\|\beta_0\|_2} = 3.1602.
\end{aligned}
$$

The largest error is e^{LS}, and the error e^{GCV} is smaller but it is very large, and much larger than the error e^{Lcurve}. The error e^{Lcurve} is smaller than the error e^{GCV} because the filter factors for $\lambda_{\text{opt}}^{\text{Lcurve}}$ are more effective than the filter factors for $\lambda_{\text{opt}}^{\text{GCV}}$ in removing the contribution of the small singular values of X to $\beta(0)$.

Figure 7 shows the variation of $\log_{10} |d_i| \, \sigma_i$, $\log_{10} |d_i + \delta d_i| \, \sigma_i$, $\log_{10} {}^{|c_i|}/_{\sigma_i}$ and $\log_{10} {}^{|c_i + \delta c_i|}/_{\sigma_i}$ with i. It is seen that $|d_i| \, \sigma_i$ and ${}^{|c_i|}/_{\sigma_i}$ are ill conditioned because they are sensitive to perturbations δd_i and δc_i, respectively, as stated in Theorem 4.1. The forward problem is well conditioned and its effective condition number is ill conditioned, but the inverse problem is different because it, and its effective condition number, are ill conditioned. This result is stated in the entry in the second row in Table 1.

Regularisation in the 1-norm (the lasso) was applied by specifying $\lambda_1 \neq 0$ and $\lambda_2 = 0$ in (4). Figure 8 shows the cross-validated mean square error using 10-fold cross-validation, and the error bars. The optimal value of the regularisation parameter and the error in the regularised solution are, respectively,

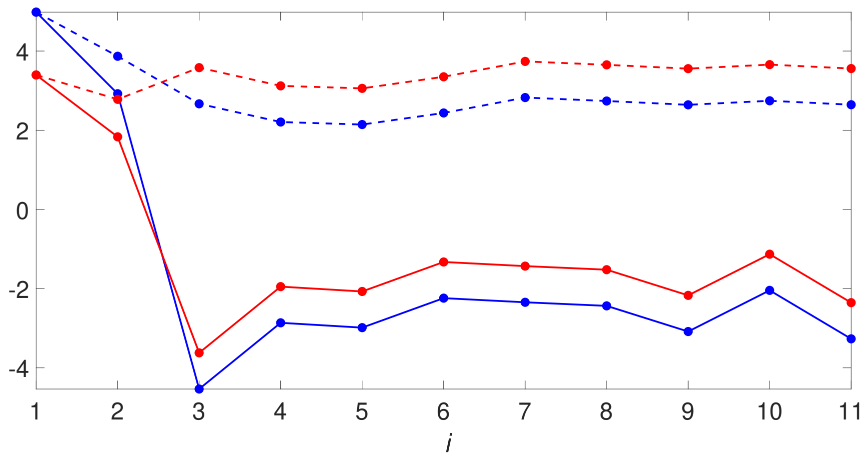

Fig. 7. The variation of $\log_{10} |d_i|\, \sigma_i$ —, $\log_{10} |d_i + \delta d_i|\, \sigma_i$ – –, $\log_{10} {}^{|c_i|}/_{\sigma_i}$ — and $\log_{10} {}^{|c_i+\delta c_i|}/_{\sigma_i}$ – – with i, for Example 6.1.

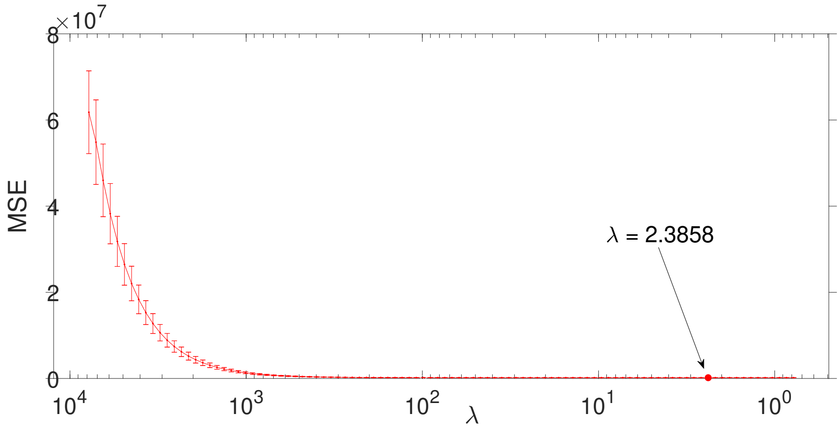

Fig. 8. The cross-validated mean square error from the lasso against the regularisation parameter λ, the error bars and the optimal value of the regularisation parameter, $\lambda = \lambda_{\text{opt}}^{\text{lasso}} = 2.3858$, for Example 6.1.

$$\lambda_{\text{opt}}^{\text{lasso}} = 2.3858 \qquad \text{and} \qquad e^{\text{lasso}} = \frac{\left\| \beta(\lambda_{\text{opt}}^{\text{lasso}}) - \beta_0 \right\|_2}{\|\beta_0\|_2} = 1.5352,$$

and the figure shows that $\lambda_{\text{opt}}^{\text{lasso}}$ is badly defined because the minimum of the curve is shallow, and it is therefore similar to Figure 5. Furthermore, the error is much larger than the errors e^{Lcurve} and e^{GCV} obtained from Tikhonov regularisation. □

Example 6.2 Consider the design matrix X in Example 6.1, but the vector β_0 is such that the forward problem is ill conditioned and the inverse problem is well conditioned,

$$\begin{aligned} \eta_1^{\text{fwd}}(X,\beta_0) &= 335.7, \qquad & \eta_1^{\text{inv}}(X,y_0) &= 3.779, \\ \eta_2^{\text{fwd}}(X,\beta_0) &= 316.2, \qquad & \eta_2^{\text{inv}}(X,y_0) &= 1.000, \end{aligned} \tag{41}$$

where β_0 and y_0 are defined in (40). The entries of ϵ are uniformly distributed random variables, and ϵ is scaled such that the SNR is equal to 10. Figure 9 shows the variation of $\log_{10}|d_i|\,\sigma_i$, $\log_{10}{}^{|c_i|}/_{\sigma_i}$, and $\log_{10}\sigma_i$ with i, and it is seen that $|d_i|\,\sigma_i$ and ${}^{|c_i|}/_{\sigma_i}$ are dominated by the small singular values σ_i, $i = 3, \ldots, 11$, of X. This must be compared with these quantities in Figure 3 for Example 6.1, which are dominated by the large singular values σ_1 and σ_2.

Figures 10 and 11 show, respectively, the L-curve and the function $G(\lambda)$ for the evaluation of the GCV, and it is seen that they differ from their equivalents in Figures 4 and 5 for Example 6.1. This difference arises because the use of the L-curve and GCV for the determination of the optimal value of λ requires that the discrete Picard condition be satisfied, but this condition is not satisfied in this example because (41) shows that the inverse problem is well conditioned. It is interesting to note that the L-curve in Figure 10 possesses a point of maximum curvature, and it is well defined. The value of λ at this point is, however, spurious as a regularisation parameter because the inverse problem is well conditioned and thus the discrete Picard condition is not satisfied. It follows that the error e^{LS} of the solution β_{LS} of the inverse problem is small,

$$e^{\text{LS}} = \frac{\|\beta_{\text{LS}} - \beta_0\|_2}{\|\beta_0\|_2} = 1.1668 \times 10^{-2}.$$

The variation of $\log_{10}|d_i|\,\sigma_i$, $\log_{10}|d_i + \delta d_i|\,\sigma_i$, $\log_{10}{}^{|c_i|}/_{\sigma_i}$ and $\log_{10}{}^{|c_i+\delta c_i|}/_{\sigma_i}$ with i is shown in Figure 12. It is clear that $|d_i|\,\sigma_i$ and ${}^{|c_i|}/_{\sigma_i}$ are stable with respect to perturbations δd_i and δc_i, respectively, which confirms Theorem 4.2 and the entry in the first row in Table 1.

Figure 13 shows the cross-validated mean square error from the lasso, and the error bars, and it is seen that the graph does not possess a minimum in the range of λ defined by the horizontal axis. $\square$

Example 6.3 Consider the design matrix X in Examples 6.1 and 6.2, but the vector β_0 is such that four of its components are equal to zero, $\beta_{0,2} = \beta_{0,3} = \beta_{0,6} = \beta_{0,10} = 0$, as shown in Figure 14. The variance of the entries in X is $\sigma^2 = 1$, the correlation coefficient is $r = 0.9999$ and the SNR is equal to 10. The effective condition numbers of the forward and inverse problems are

$$\begin{aligned} \eta_1^{\text{fwd}}(X,\beta_0) &= 1.064, \qquad & \eta_1^{\text{inv}}(X,y_0) &= 1192, \\ \eta_2^{\text{fwd}}(X,\beta_0) &= 3.149, \qquad & \eta_2^{\text{inv}}(X,y_0) &= 100.4, \end{aligned}$$

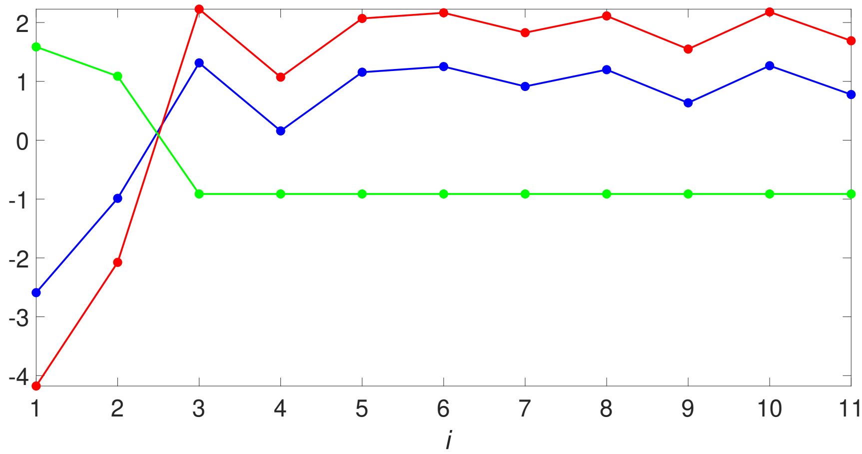

Fig. 9. The variation of $\log_{10} |d_i| \, \sigma_i$ ●, $\log_{10} {}^{|c_i|}/_{\sigma_i}$ ● and $\log_{10} \sigma_i$ ● with i, for Example 6.2.

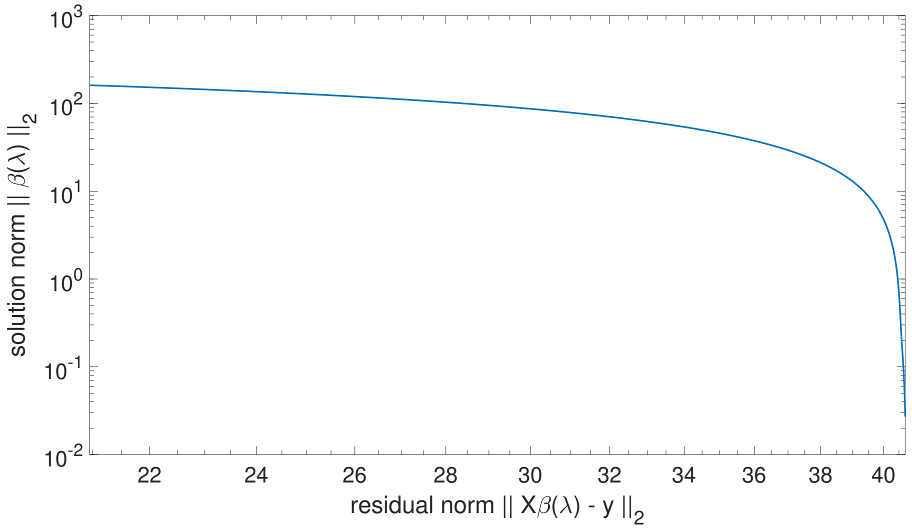

Fig. 10. The L-curve for Example 6.2.

and the terms $|d_i| \, \sigma_i$ and ${}^{|c_i|}/_{\sigma_i}$ are sensitive to perturbations δd_i and δc_i, respectively, because the forward problem is well conditioned and the inverse problem is ill conditioned, as shown in Figure 7 for Example 6.1.

Tikhonov regularisation and the lasso were used to regularise the inverse problem, and the optimal values of the regularisation parameters were

$$\lambda_{\text{opt}}^{\text{Lcurve}} = 0.3186, \qquad \lambda_{\text{opt}}^{\text{GCV}} = 0.03227, \qquad \lambda_{\text{opt}}^{\text{lasso}} = 7.523 \times 10^{-5},$$

and thus $\lambda_{\text{opt}}^{\text{lasso}} \ll \lambda_{\text{opt}}^{\text{GCV}}, \lambda_{\text{opt}}^{\text{Lcurve}}$. The vector $\beta(\lambda_{\text{opt}}^{\text{Lcurve}})$ is shown in Figure 15, and the vector $\beta(\lambda_{\text{opt}}^{\text{GCV}})$ is very similar because its first component is large and dominant, and its other components are non-zero and much smaller. There is

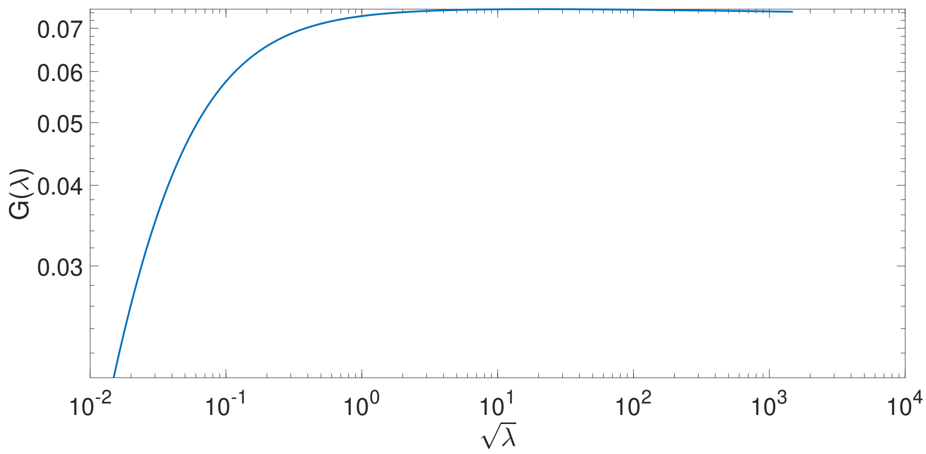

Fig. 11. The GCV function $G(\lambda)$ against $\sqrt{\lambda}$ for Example 6.2.

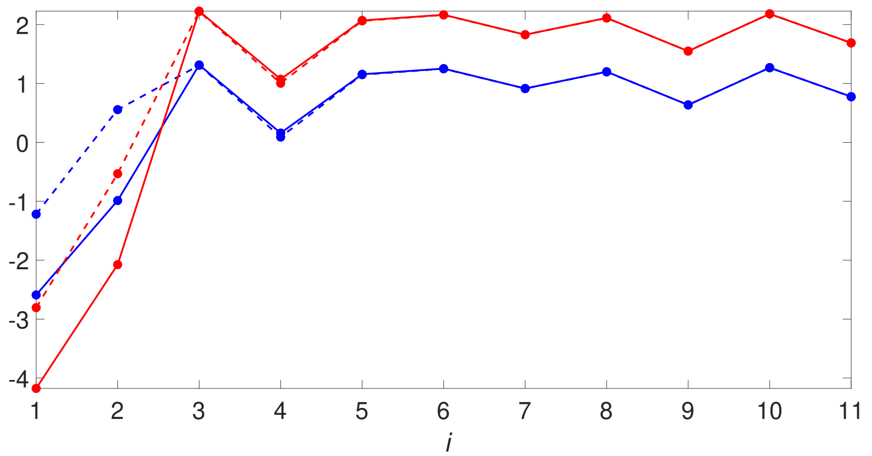

Fig. 12. The variation of $\log_{10} |d_i|\, \sigma_i$ —, $\log_{10} |d_i + \delta d_i|\, \sigma_i$ − −, $\log_{10} {}^{|c_i|}/_{\sigma_i}$ — and $\log_{10} {}^{|c_i+\delta c_i|}/_{\sigma_i}$ − − with i, for Example 6.2.

therefore a clear distinction between the entries of the solutions from Tikhonov regularisation that are non-zero and dominant, and the entries that are significantly smaller and are therefore approximately zero. The solution $\beta(\lambda_{\text{opt}}^{\text{lasso}})$ from the lasso was unsatisfactory because many components of β_0 that are either equal to zero or small were much larger in $\beta(\lambda_{\text{opt}}^{\text{lasso}})$. The relative errors of the LS solution β_{LS}, the solutions from Tikhonov regularisation and the lasso are

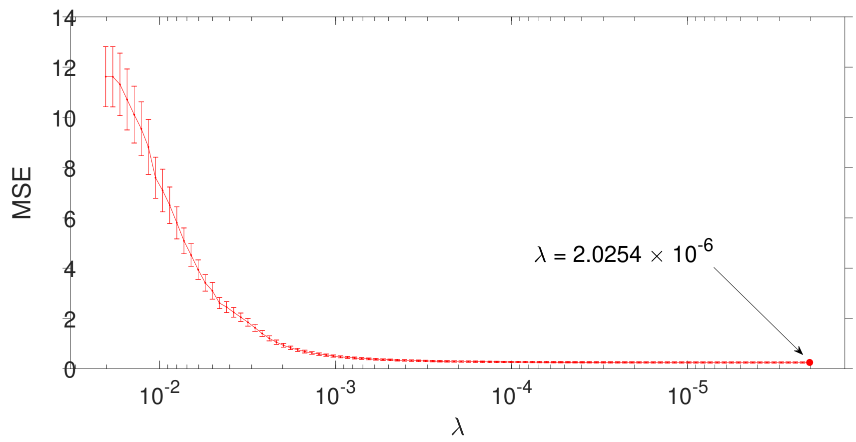

Fig. 13. The cross-validated mean square error from the lasso, and the error bars, for Example 6.2.

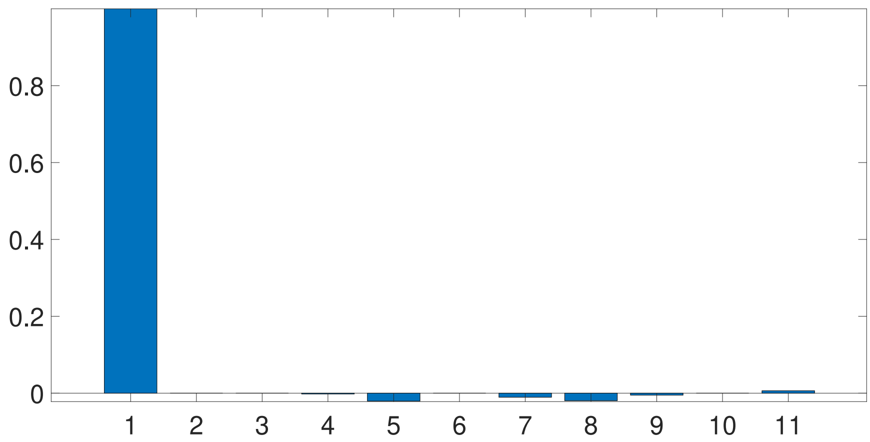

Fig. 14. The vector β_0 for Example 6.3.

$$e^{\mathrm{LS}} = \frac{\|\beta_{\mathrm{LS}} - \beta_0\|_2}{\|\beta_0\|_2} = 1.1906,$$

$$e^{\mathrm{Lcurve}} = \frac{\left\|\beta(\lambda_{\mathrm{opt}}^{\mathrm{Lcurve}}) - \beta_0\right\|_2}{\|\beta_0\|_2} = 0.1027,$$

$$e^{\mathrm{GCV}} = \frac{\left\|\beta(\lambda_{\mathrm{opt}}^{\mathrm{GCV}}) - \beta_0\right\|_2}{\|\beta_0\|_2} = 0.3852,$$

$$e^{\mathrm{lasso}} = \frac{\left\|\beta(\lambda_{\mathrm{opt}}^{\mathrm{lasso}}) - \beta_0\right\|_2}{\|\beta_0\|_2} = 1.000,$$

and thus the error in $\beta(\lambda_{\mathrm{opt}}^{\mathrm{lasso}})$ is much larger than the errors in $\beta(\lambda_{\mathrm{opt}}^{\mathrm{Lcurve}})$ and $\beta(\lambda_{\mathrm{opt}}^{\mathrm{GCV}})$. The error e^{Lcurve} is smaller than the error e^{GCV} because the filter factors for $\lambda_{\mathrm{opt}}^{\mathrm{Lcurve}}$ are more effective than the filter factors for $\lambda_{\mathrm{opt}}^{\mathrm{GCV}}$ in removing the contribution of the small singular values of X to $\beta(0)$. The large

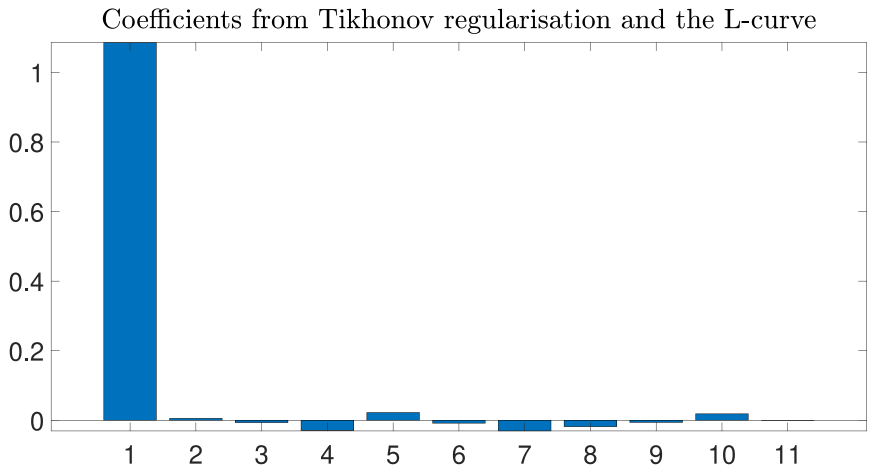

Fig. 15. The solution vector $\beta(\lambda_{\text{opt}}^{\text{Lcurve}})$ for Example 6.3.

error in β_{LS} is expected because the inverse problem is ill conditioned. □

7 Summary

This paper has considered numerical issues in regularisation and variable selection when X is a design matrix with correlated covariates. It has been shown that the stability of the forward problem and the stability of the inverse problem are not independent because the product of their effective condition numbers is equal to the condition number of X. Also, Tikhonov regularisation requires that the discrete Picard condition be satisfied because this guarantees that the regularisation error is small and the regularised solution is stable, provided that the singular values of X decay to zero.

It was shown that $\tilde{X}$, where $\tilde{X}^T\tilde{X} = E\left\{X^TX\right\}$, has two distinct singular values and one multiple singular value, and formulae for them were derived. It was shown that the filter factors and singular values of $\tilde{X}$ do not decay to zero but level off at a constant value.

The lasso was investigated experimentally but the results were unsatisfactory because the regularisation error was large, even for ill conditioned problems. It is therefore necessary to investigate the properties of the lasso, and in particular, the conditions to be satisfied for it to yield an acceptable solution must be determined.

The work described in this paper can be extended to other design matrices, for example, matrices with correlated columns but different correlation coefficients. The decay of the singular values of these design matrices, and the discrete Picard condition, must be considered because they are important in determining the success, or otherwise, of the application of Tikhonov regularisation to the inverse problem in variable selection.

References

[1] Craven, P., Wahba, G.: Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. Numerische Mathematik **31**(4), 337–403 (1978)

[2] Forman, G.: An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. **3**, 1289–1305 (2003)

[3] Golub, G.H., Heath, M., Wahba, G.: Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics **21**(2), 215–223 (1979)

[4] Golub, G.H., Meurant, G.: Matrices, Moments and Quadrature with Applications. Princeton University Press, Princeton, USA (2009)

[5] Hansen, P.C.: Analysis of discrete ill-posed problems by means of the L-curve. SIAM Review **34**(4), 561–580 (1992)

[6] Hansen, P.C.: Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion. SIAM, Philadelphia, USA (1998)

[7] Hansen, P.C.: Regularization Tools version 4.0 for Matlab 7.3. Numer. Algor. **46**(2), 189–194 (2007)

[8] Hansen, P.C., Nagy, J.G., O'Leary, D.P.: Deblurring Images: Matrices, Spectra, and Filtering. SIAM, Philadelphia, USA (2006)

[9] Koukouvinos, C., Mitrouli, M., Turek, O.: Efficient estimates in regression models with highly correlated covariates. J. Comp. Appl. Math. **373** 112416 (2020)

[10] Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Statis. Soc. B **58**(1), 267–288 (1996)

[11] Wahba, G.: Practical approximate solutions to linear operator equations when the data are noisy. SIAM J. Numer. Anal. **14**, 651–667 (1977)

[12] Wang, H., Nie, F., Huang, H., Risacher, S., Ding, C., Saykin, A., Shen, L.: Sparse multi-task regression and feature selection to identify brain imaging predictors for memory performance. In: Proc. IEEE Int. Conf. Computer Vision, pp. 557–562. Barcelona, Spain (2011)

[13] Winkler, J.R., Mitrouli, M.: Condition estimation for regression and feature selection. J. Comp. Appl. Math. **373** 112212 (2020)

[14] Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. J. R. Statist. Soc. B **67**(2), 301–320 (2005)