Chapter 1

Inversion of 2D NMR Data

Christopher Bose[1], Zhenlu Cui[2], Xinghua Deng[3], Ying Han[4], Qingguo Li[5], Robert Piché[6], Lalitha Venkataramanan[7], Qian Wang[3], Lin Zhou[8]

Report prepared by Christopher Bose (cbose@math.uvic.ca)

1.1 Introduction

Schlumberger Limited is a multinational company supplying oilfield and information services to a worldwide energy market. These services include both exploration and production tools ranging through seismic and remote sensing, well-logging and reservoir optimization. The problem described in this report is related to well-logging via Nuclear Magnetic Resonance (NMR), a relatively new and developing tool with potential to reveal a range of reservoir properties including porosity and saturation, as well as physical properties of the petroleum deposit.

In order to recover this information from NMR spectra the company must have an effective, efficient and robust algorithm to perform inversion from the dataset to the unknown probability distribution on magnetic relaxation times. This ill-posed problem is encountered in diverse areas of magnetic imaging and there does not appear to be an 'off-the-shelf' solution which the company can apply to its problem. Company scientists have developed a sophisticated algorithm which performs well on some simple test datasets, but they are interested in knowing if there are simpler approaches which could work effectively, or if some limited but useful properties of the density are accessible with a totally different approach.

Our report is organised as follows. In Section 1.2 we present a careful and complete description of the problem and the work already done by the company. In Section 1.3 we discuss Truncated Singular Value Regularisation and Tikhonov Regularisation and show how some 'off-the-shelf' Matlab code

[1]University of Victoria
[2]Florida State University
[3]University of Alberta
[4]McGill University
[5]Simon Fraser University
[6]Tampere University of Technology, Finland
[7]Schlumberger Doll Research
[8]New Jersey Institute of Technology

may be used to good effect on the test datasets provided by the company. In Section 1.4 we show that one can incorporate higher order regularisation into the company's existing algorithm, answering one specific question raised at the beginning of the workshop. Finally, in Section 1.5 we record our unsuccessful attempt to establish an iterative algorithm for the positively constrained inversion. Finally in the last section we review our conclusions and make suggestions for future work.

1.2 Problem Description

Schlumberger is interested in using Nuclear Magnetic Resonance (NMR) analysis for exploration in the oil and gas industry. The model problem presented to our group at the workshop involved the recovery of a two-dimensional probability distribution $f(x, y)$ on magnetic field relaxation times in two directions, x, the so-called longitudinal relaxation time and y the transverse relaxation time. The data collected is known to be a convolved image of the relaxation time distribution according to the following formula

$$d(\tau_1, \tau_2) = \iint (1 - 2e^{-\tau_1/x})e^{-\tau_2/y} f(x, y)\, dx\, dy. \tag{1.1}$$

Other types of data can be collected, involving different convolution kernels, but the forward model, in any case is in the form of a 2-D **Fredholm Integral of the First Kind**. Since the transformation involves a smooth kernel, it is well known that the corresponding inverse problem is ill-posed [5, p.2]. None the less, it is important for the company's program to provide some sort of stable and computationally tractable inversion scheme.

The continuous forward model (1.1) is mainly of theoretical interest since in practice, data is collected at discrete values in the $\tau_1\tau_2$-domain. Therefore, for the rest of this analysis we will assume the data function d is replaced by a data **matrix** D of dimension $m_2 \times m_1$. Convolution kernels are similarly discretized as matrices K_1 and K_2 with dimensions $m_1 \times n_1$ and $m_2 \times n_2$ respectively and a discrete form of the equation (1.1) is rewritten as

$$D = K_2 F K_1^T. \tag{1.2}$$

The discrete density F is now an $n_2 \times n_1$ matrix. K_1 and K_2 are (generally) rank-deficient with infinite condition number and singular values decaying quickly to zero. So, as expected, the ill-posed problem leads to an ill-conditioned finite-dimensional inversion (1.2).

Three sets of test data were provided to our group for use during the workshop. Distribution files F were of size $n_2 \times n_1 = 100 \times 100$. Kernel discretisation led to K_1 and K_2 of size 30×100 and 4000×100 respectively. There are reasonable grounds for the asymmetric choice in the discretisation grid here. For test inversion problems we replace the data D computed from equation (1.2) by $\hat{D} = D + E = K_2 F K_1^T + E$ where E is mean-zero Gaussian noise. The object is to recover F. The choice of signal to noise ratio for the various test files will be discussed later in the numerical results section.

Suppose for the moment we take a completely naive point of view and convert our problem to a standard one-dimensional least squares approximation

$$\text{vec}(\hat{F}) = \arg\min_f \|Kf - \text{vec}(\hat{D})\|^2 \tag{1.3}$$

where vec$(\cdot)$ represents the operator making a vector from a matrix by stacking columns, $K = K_1 \otimes K_2$ is the Kronecker product of the convolution kernels and $f = \text{vec}(F)$. Since K is huge ($m_1 m_2 \times n_1 n_2 = 120,000 \times 10,000$) and dense, we may have difficulty fitting it into the RAM memory of a PC even if we ignore the computational complexity of the positivity constraint $f \geq 0$ and ill-posed nature of the high dimension inversion! For example, in [3] a similar problem arising in medical imaging was analysed using a CRAY supercomputer. Therefore, we conclude that a numerically reasonable approach of the type required by the company should try to work directly with the factored form (1.2). This observation was known to the company scientists. For this reason, most of our analysis will be centred on the factored problem of the type

$$\hat{F} = \arg\min_{F \geq 0} \|K_2 F K_1^T - \hat{D}\|_{\text{Fro}}^2 \tag{1.4}$$

where $\|\cdot\|_{\text{Fro}}$ denotes the Frobenius matrix norm.

The problem proposor (L.V.) described a three step approach to the the optimisation in (1.4) which the company has found to be effective on the test datasets. First, the problem dimension is significantly reduced by projection. The range of this projection is related to the singular value decomposition (SVD) truncation of the convolution matrices. Next, the ill-conditioned (but lower-dimension) problem is regularised as a positively constrained Tikhonov optimisation in unfactored form

$$\text{vec}(\hat{F}) = \arg\min_{f} \left(\|Kf - \text{vec}(\hat{D})\|^2 + \lambda^2 \|f\|^2 \right) \tag{1.5}$$

where λ is the regularisation parameter. Finally, this constrained problem is solved by the method in Butler, Reeds and Dawson [1] (BRD) which transforms to an unconstrained optimisation with respect to a derived objective function. Details, including methods to choose the regularisation parameter and performance on the test problems may be found in [10].

With this background in place our group was asked to consider three lines of investigation.

First, are there other numerically tractable (and possibly simpler) approaches to the inversion problem (1.2)? In the next section, we describe three answers to this question. First we consider using truncated singular value decomposition (TSVD) and Tikhonov regularisation on the factored form (1.4), greatly reducing the computational and algorithmic complexity of previous methods. Performance on the test datasets is presented. We also consider briefly a direct Galerkin-type approach.

Next, we were asked consider the possibility of extending the BRD-method described by the proposor to higher-order Tikhonov regularization. In particular, can we replace the problem (1.5) with

$$\hat{F} = \arg\min_{F \geq 0} \left\{ \|K_2 F K_1^T - \hat{D}\|_{\text{Fro}}^2 + \lambda^2 (\|LF\|_{\text{Fro}}^2 + \|FL^T\|_{\text{Fro}}^2) \right\} \tag{1.6}$$

where L invokes the discrete first derivatives on the square matrix F? We present mixed results for this second question in that we can transform the problem (1.6) into a standard problem of the type (1.5), to which the BRD method can subsequently be applied, but we cannot arrange that the transformed problem has the desirable Kronecker product structure. In a slightly different direction we consider if the company's idea for an iterative algorithm can be adapted to higher-order regularisation. Unfortunately the same problems which led to the use of the BRD method appear to confound this approach as well.

Finally, the company scientists believe that it may not be necessary to obtain complete inversion of the problem but that some macroscopic information about the distribution of relaxation times (moments or (x_1, x_2)-correlations for example) may be sufficient. While this question may be amenable to a Galerkin approach, without prior information about a restricted class of possible distributions our group saw no tractable way to make progress in this direction during the week of the workshop.

1.3 TSVD and Tikhonov Regularisation

The Theory

Let us first set up a unified framework for these two well known regularisation methods in non-factored problems. Consider the discrete linear system

$$Kf = d \tag{1.7}$$

where K is $m \times n$. Let $K = U\Sigma V^T$ be a SVD where Σ is $m \times n$ diagonal, U is $m \times m$ orthogonal, and V is $n \times n$ orthogonal. The Tikhonov and TSVD regularised solutions of (1.7) are

$$f_{\text{reg}} = V\phi(\Sigma^T)U^T d \tag{1.8}$$

where

$$\underset{\text{Tikhonov}}{\phi(\sigma) = \frac{\sigma}{\sigma^2 + \lambda^2}}, \qquad \underset{\text{TSVD}}{\phi(\sigma) = \begin{cases} 1/\sigma & \text{if } |\sigma| > \lambda \\ 0 & \text{otherwise} \end{cases}}$$

is applied elementwise. The nonnegative regularisation parameter λ affects the amount of smoothing of the regularised solution. Its value can be selected by minimising the generalised cross validation (GCV) function

$$G(\lambda) = \left(\frac{\|Kf_{\text{reg}} - d\|}{\text{trace}(U\Sigma\phi(\Sigma^T)U^T - I)} \right)^2 = \left(\frac{\|c \circ (U^T d)\|}{\text{sum}(c)} \right)^2$$

where the vector norm is Euclidean, $\circ$ is the Hadamard product (elementwise multiplication), and $c = \text{diag}(\Sigma\phi(\Sigma^T)) - 1$.

Now remember that the coefficient matrix in (1.7) has the Kronecker product structure

$$K = K_1 \otimes K_2$$

where K_1 is $m_1 \times n_1$ and K_2 is $m_2 \times n_2$ with $m_1 m_2 \geq n_1 n_2$. As we have observed, the key to effective algorithms is to rewrite formulae in ways that avoid explicitly forming the Kronecker product of full matrices.

Let $F = \text{reshape}_{n_2 \times n_1}[f]$ and $D = \text{reshape}_{m_2 \times m_1}[d]$. Then the linear system (1.7) is obtained by applying the vec operator to both sides of the matrix equation

$$K_2 F K_1^T = D \tag{1.9}$$

This equation does not involve the Kronecker product. Similar techniques can be used to eliminate expensive Kronecker products from the regularisation formulae, as follows.

Let $K_1 = U_1\Sigma_1 V_1^T$ and $K_2 = U_2\Sigma_2 V_2^T$ be SVDs. Then $(U_1 \otimes U_2)(\Sigma_1 \otimes \Sigma_2)(V_1 \otimes V_2)^T$ is an SVD of K [7, Thm 4.2.15]. The formula (1.8) for the regularised solution can therefore be written

$$\begin{aligned} F_{\text{reg}} &= \text{reshape}_{n_2 \times n_1}[f_{\text{reg}}] \\ &= \text{reshape}_{n_2 \times n_1}[(V_1 \otimes V_2)\phi(\Sigma_1^T \otimes \Sigma_2^T)(U_1 \otimes U_2)^T d] \\ &= V_2 \cdot \text{reshape}_{n_2 \times n_1}[\phi(\Sigma_1^T \otimes \Sigma_2^T)\text{vec}(U_2^T D U_1)] \cdot V_1^T. \end{aligned} \tag{1.10}$$

The Kronecker product formation and multiplication in (1.10) only involves diagonal matrices, so the formula can be implemented efficiently with appropriate data structures. Further savings are possible by using the "economy size" version of the SVD of K_2 when $m_2 > n_2$.

Similarly, the formula for the GCV function can be written

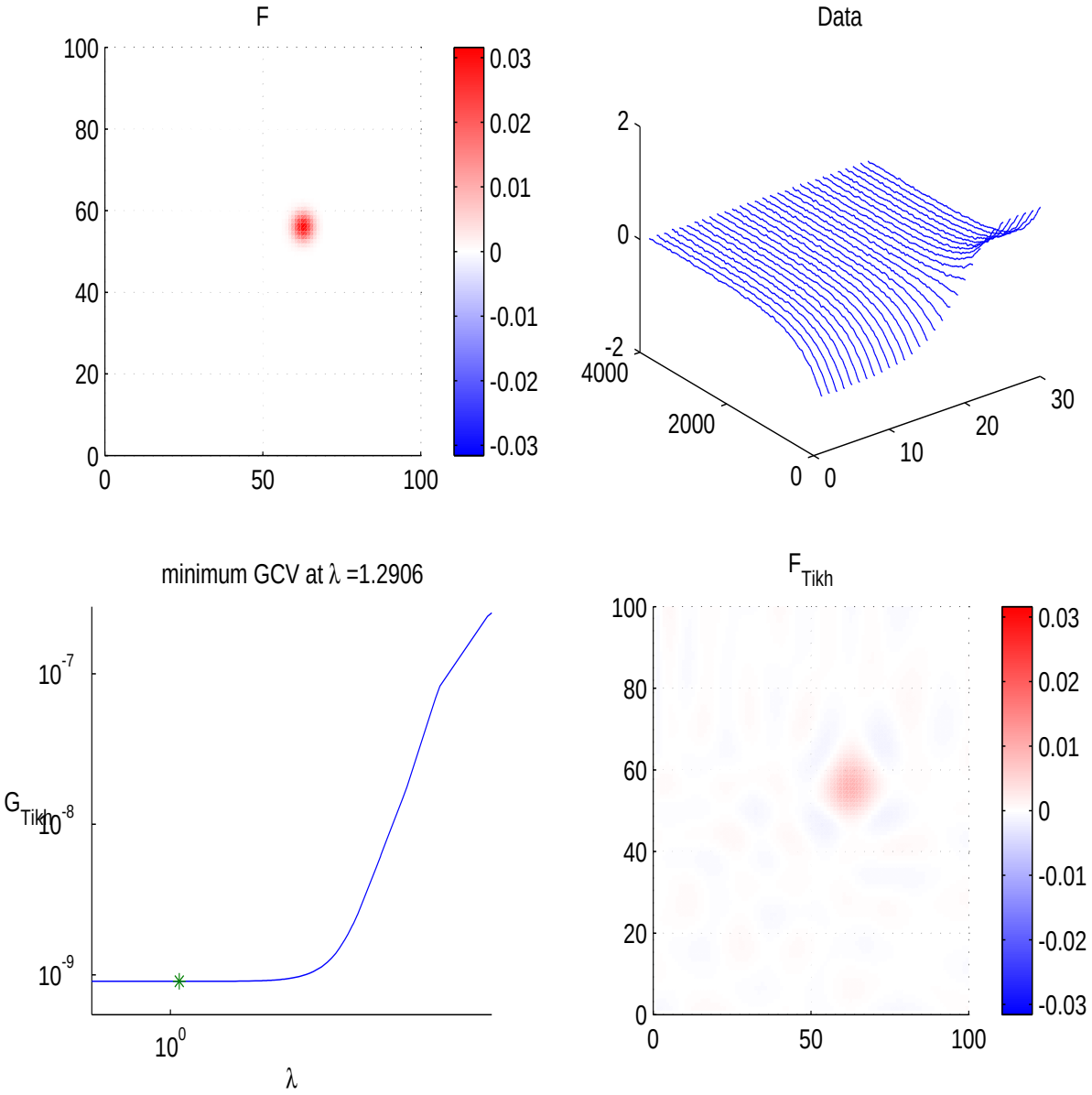
$$G(\lambda) = \left(\frac{\|C \circ (U_2^T D U_1)\|_{\text{Fro}}}{\text{sum}(C)} \right)^2$$

where $C = \text{reshape}_{m_2 \times m_1}[\text{diag}((\Sigma_1 \otimes \Sigma_2)\phi(\Sigma_1^T \otimes \Sigma_2^T))] - 1$. Here again the only Kronecker products are of diagonal matrices.

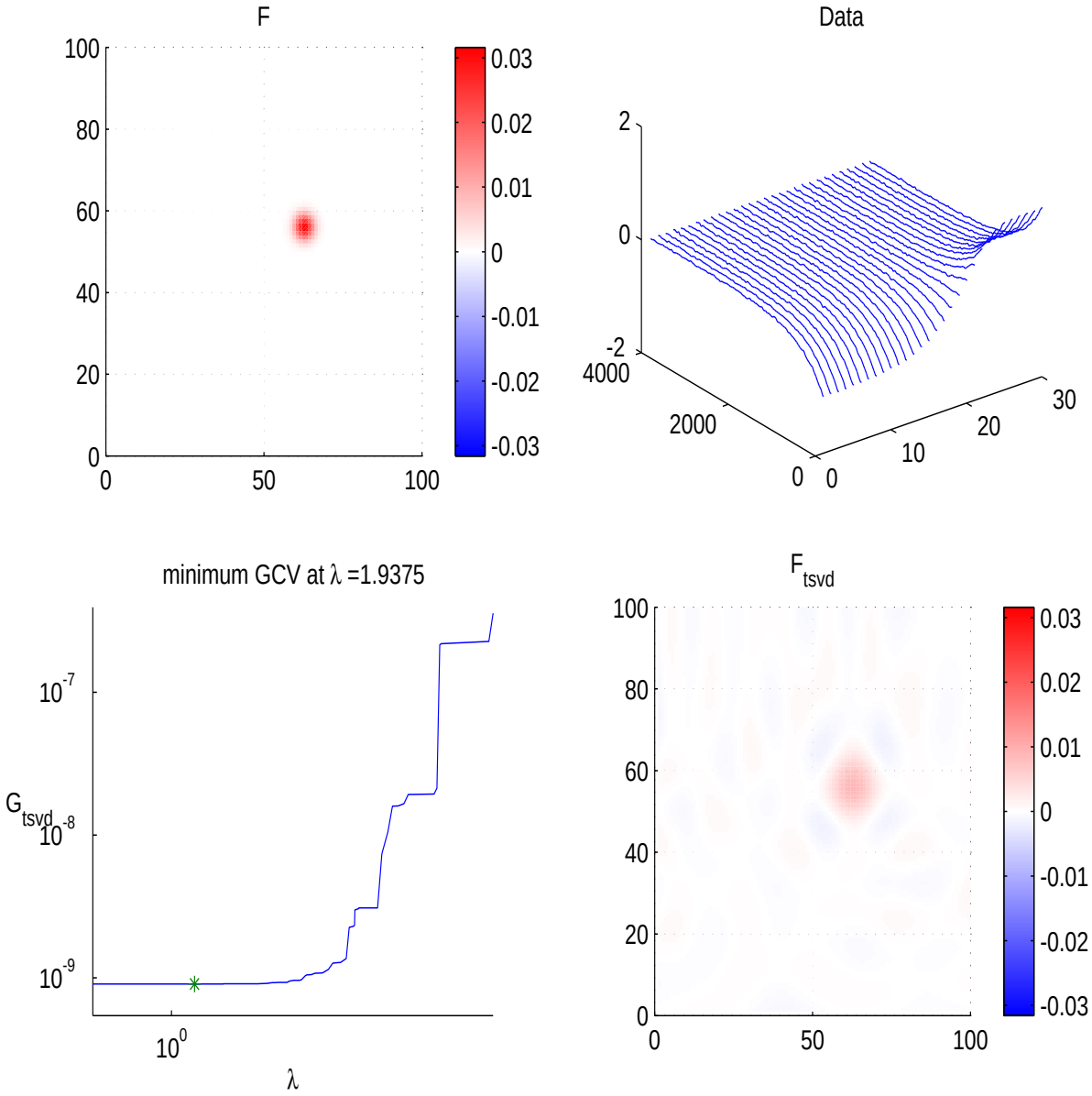
Numerical Results

The problem proposer provided K_1 (of size 30×100), K_2 (of size 4000×100), and three different 100×100 F matrices. Measurement data was generated by adding zero mean pseudorandom noise E to $K_2 F K_1^T$; the noise variance was set so that $\|E\|_{\text{Fro}} = 0.05\|K_2 F K_1^T\|_{\text{Fro}}$.

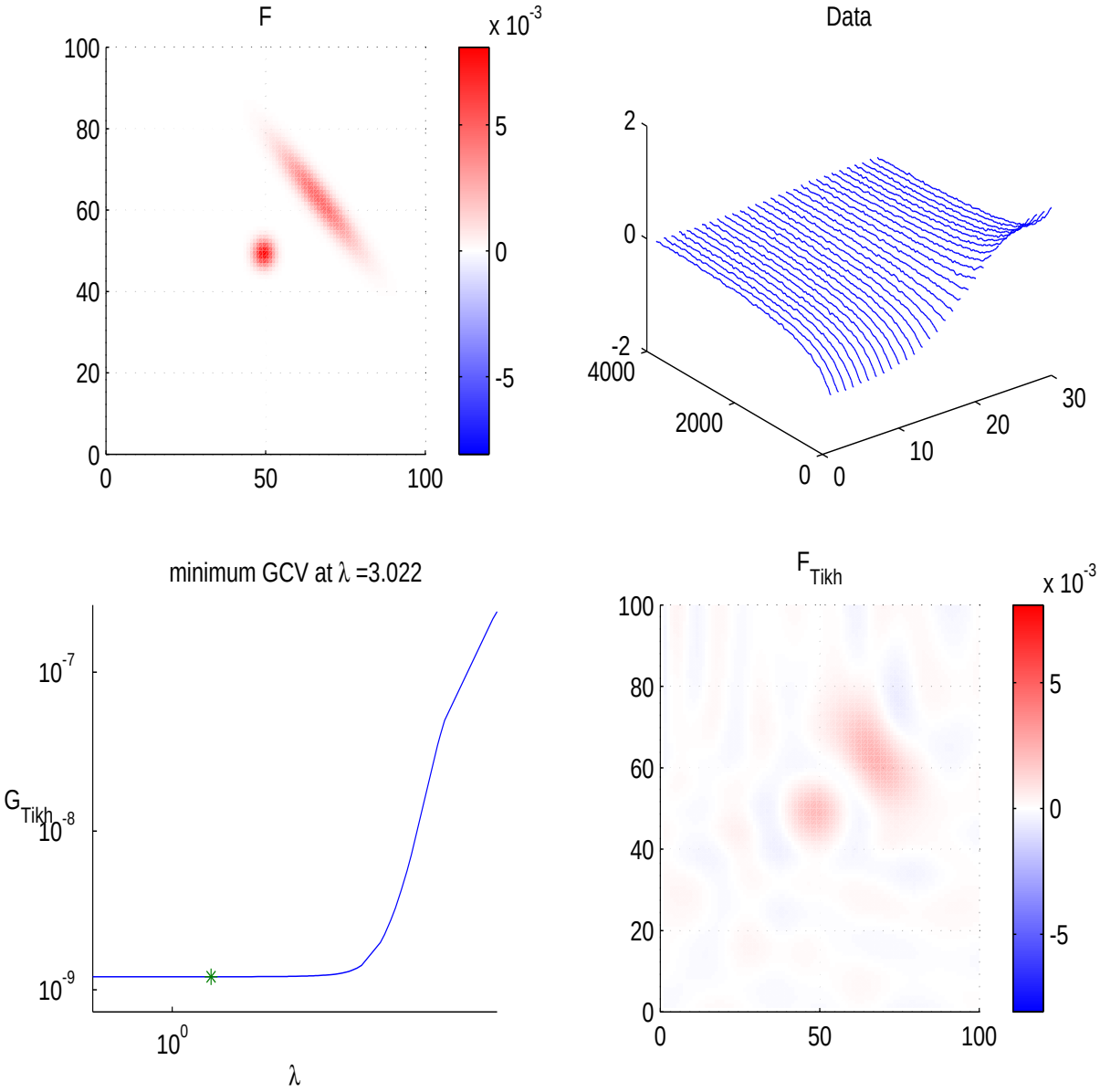
Each of the following regularisations (including SVD and GCV curve computations) took about 12 seconds to compute in Matlab 5.2 on a 30 MB memory partition of a 400 MHz Powerbook. The GCV minimisation appears to select reasonable regularisation parameters, and Tikhonov and TSVD regularisation give about the same results for all three models. The data and Matlab code are available at `http://alpha.cc.tut.fi/~piche/ipsw2003/`



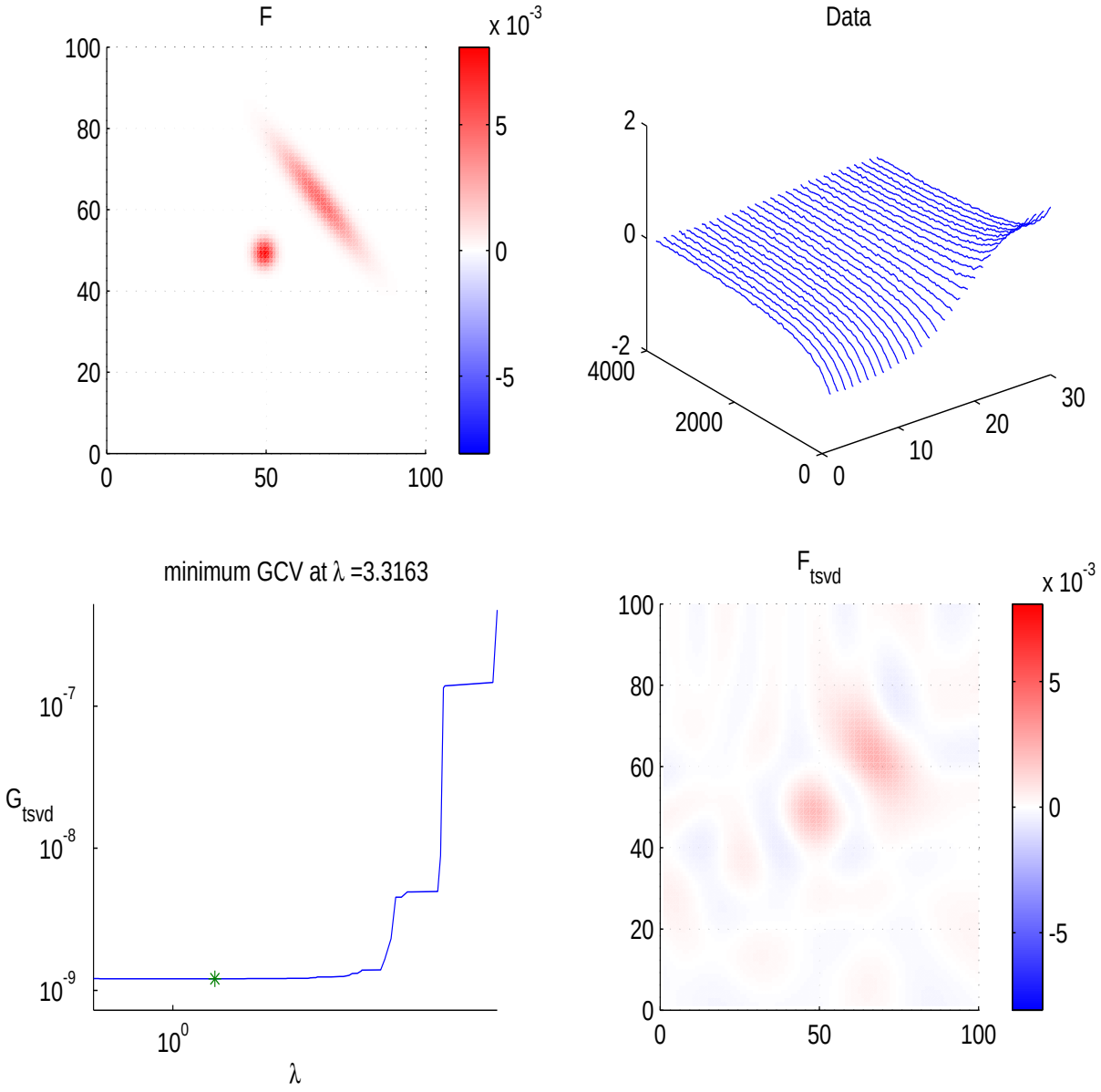
model A, Tikhonov regularisation



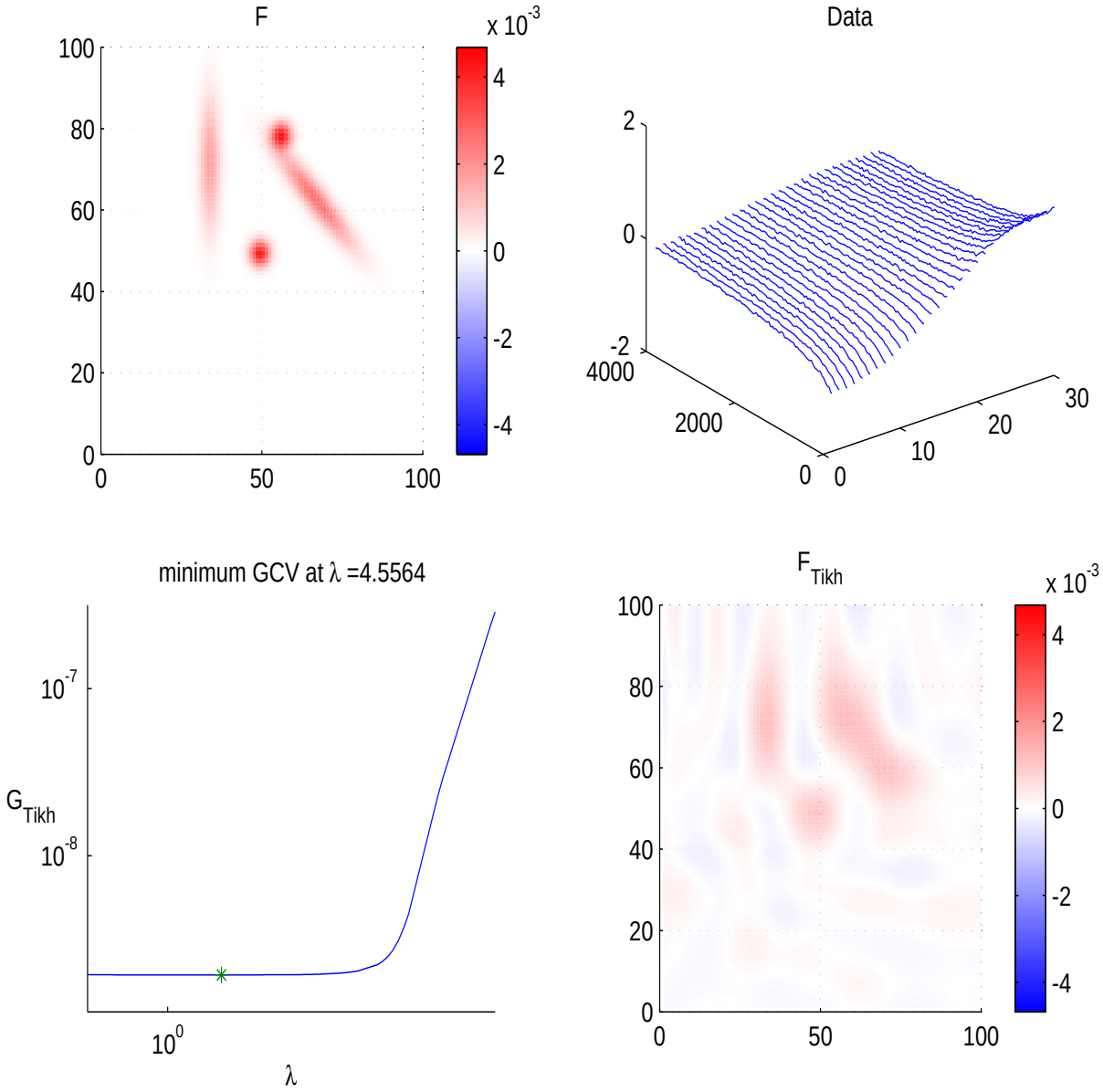
model A, TSVD regularisation



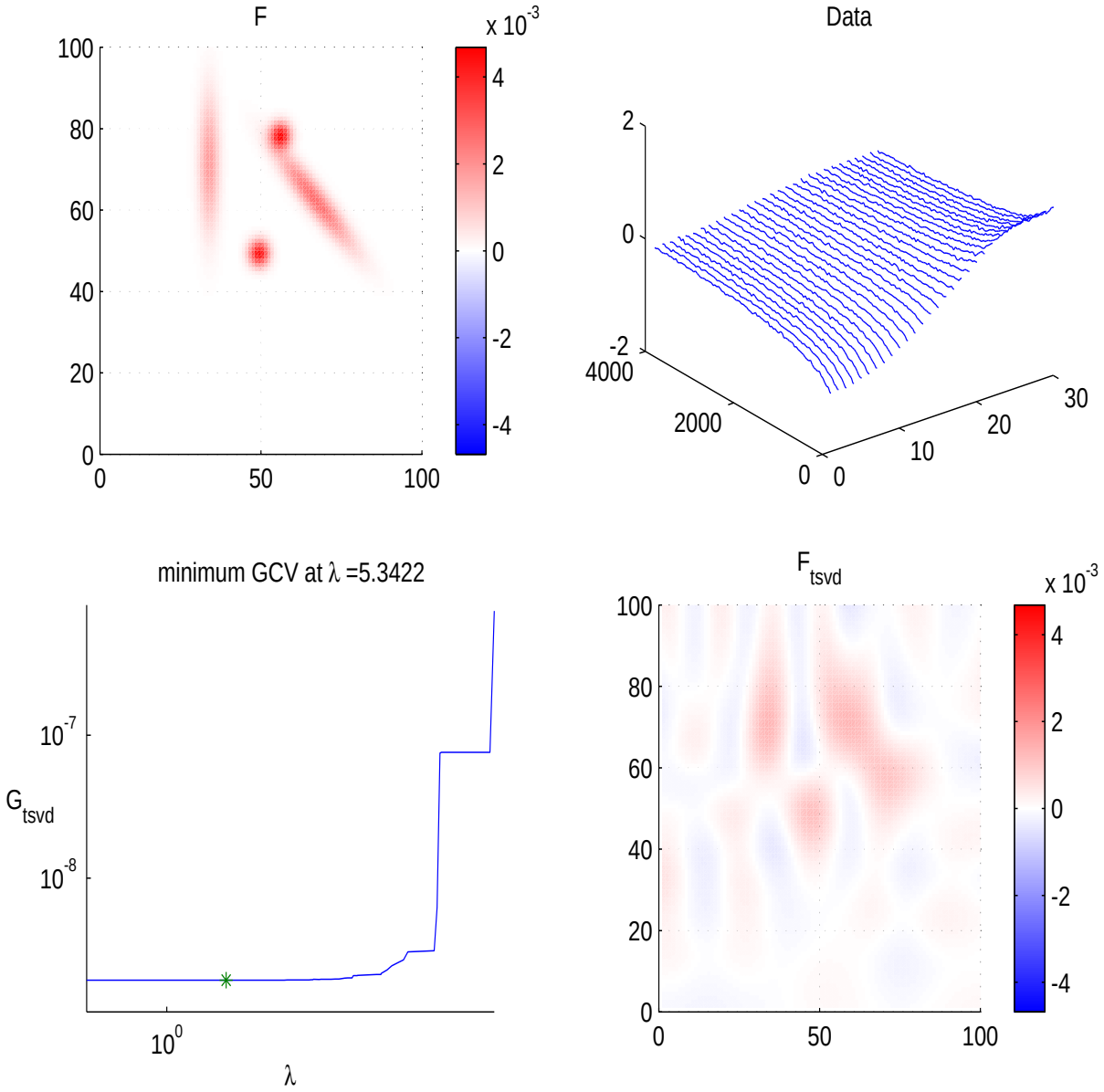
model B, Tikhonov regularisation



model B, TSVD regularisation



model C, Tikhonov regularisation



model C, TSVD regularisation

Parameterised Methods

Our group briefly considered the possibility of using a Galerkin approach to the inversion problem (1.1). Thus, we make the ansatz

$$f(x,y) = \sum_i f(x,y,\mathcal{P}_i) \tag{1.11}$$

where the $f(x,y,\mathcal{P}_i)$ are a finite set of parameterised basis functions with parameter values $\mathcal{P}_i$.

We remark that this approach reduces to the analysis of the previous paragraphs by the choice of the basis functions as 'delta functions' $f(\cdot,F_{ij}) = F_{ij}\delta_(x_i,y_j)(\cdot)$ centred on the points of the discretisation lattice and the optimal parameter selection is the matrix $\hat{F}$ of the previous analysis. As we have discovered this is a high-dimensional, ill-posed and (because of the positivity requirement) nonlinear problem. Our question then is this: Can a judicious choice of basis functions lead to a significantly smaller parameter space (instead of the 10,000 - dimensional space already encountered)? Of course the ill-posed nature of the problem must reappear in the Galerkin method as the number of basis functions increases, no matter how cleverly this basis is chosen.

The following are a few examples for the basis functions which seem well-suited to the test problems given to the group.

1. Gaussian functions ($\mathcal{P} = \{x_0, y_0, \sigma, P\}$)

$$g(x,y) = \frac{1}{\sqrt{2\pi\sigma}}\, e^{\mathbf{x}^T P^T P \mathbf{x}}$$

   where $\mathbf{x} = (x - x_0, y - y_0)^T$.

2. Box functions with centre at (x_0, y_0), and dimension of its base are $2a$ and $2b$ and height $1/4ab$.

3. Pyramid functions with centre at (x_0, y_0), and square base dimension of $2a$ and $2b$ and height $3/4ab$.

For example, when using a basis consisting of one box function ($\mathcal{P} = \{c, x_0, y_0, a, b\}$) to approximate $f(x,y)$ the Fredholm integral becomes

$$\begin{aligned} D_{\mathcal{P}}(\tau_1,\tau_2) &= \frac{c}{4ab}\int_{x_0-a}^{x_0+a}\int_{y_0-b}^{y_0+b}(1-2e^{-\tau_1/x})(e^{-\tau_2/y})\,dy\,dx \\ &= \frac{c}{4ab}\int_{x_0-a}^{x_0+a}(1-2e^{-\tau_1/x})\,dx\int_{y_0-b}^{y_0+b}e^{-\tau_2/y}\,dy. \end{aligned} \tag{1.12}$$

There is no closed form solution to this integral, which can only be solved numerically.

Based on this simple parameterization, the problem of approximating $f(x,y)$ is transformed as a nonlinear optimization problem stated as follows:

For given $\tau_1 \in [0,T_1], \tau_2 \in [0,T_2]$ and data $\hat{D}(\tau_1,\tau_2)$, the objective is to find $\mathcal{P} = \{c, x_0, y_0, a, b\}$, $c \geq 0$ such that

$$\hat{F} = \underset{f(x,y,\mathcal{P})}{\arg\min} \|D_{\mathcal{P}} - \hat{D}\|^2_{\mathrm{Fro}}.$$

Unfortunately, due to time constraints we were unable to conduct numerical tests on this optimisation problem during the week of the workshop.

We note that even though this overly simplified approach has no hope of establishing fine structure of the underlying density, it would be interesting to see if macroscopic properties desired by the company scientists could be isolated with such a relatively low-dimensional parameterisation. On the other hand, the method depends on *a priori* information about the the density, likely a fatal flaw for any robust numerical package of the type required by the company.

1.4 Higher-Order Tikhonov Regularisation

The Tikhonov regularised solution described Section 1.3 is the minimiser of the objective function

$$\frac{1}{2}\|Kf - d\|^2 + \frac{1}{2}\lambda^2\|f\|^2 = \frac{1}{2}\|K_2 F K_1^T - D\|_{\text{Fro}}^2 + \frac{1}{2}\lambda^2\|F\|_{\text{Fro}}^2. \tag{1.13}$$

A more general regularisation has the objective function's second term in the form

$$\frac{1}{2}\lambda^2\|LF\|_{\text{Fro}}^2$$

where the operator L is chosen to penalise undesired features of the solution. When L has the same Kronecker product structure as K, then it is straightforward to develop efficient regularisation algorithms along the lines of the previous section.

In this section we will consider the more difficult problem of incorporating a non-factored regularisation term. For example, setting

$$L = \begin{pmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & & -1 & 1 \\ & & & & & -1 \end{pmatrix}_{(n+1)\times n}$$

the discrete first derivative, and changing the regularisation term in (1.13) to be

$$\frac{1}{2}\lambda^2(\|LF\|_{\text{Fro}}^2 + \|FL^T\|_{\text{Fro}}^2)$$

in effect regularises by the **boundary value problem**

$$\Delta_d F = 0$$

$$F(1,j) = F(n,j) = 0 \quad \text{for all } j$$

$$F(i,1) = F(i,n) = 0 \quad \text{for all } i$$

where Δ_d denotes the discrete Laplacian. The boundary conditions ensure that L has trivial kernel which will be useful for us later. Similar considerations would allow regularisation with respect to higher order derivatives, for example replacing L with the discrete Laplacian operator plus appropriate boundary conditions to ensure a trivial kernel.

The estimation of F is equivalent to solving the following problem

$$\hat{F} = \underset{F \geq 0}{\arg\min} \left\{ \|K_2 ... etc ... \|_{\text{Fro}}^2 \right\}, \tag{1.14}$$

where K_1, K_2 are the convolution kernels and $\hat{D}$ is noisy data. Note that in this section we are keeping the notation simple by assuming a square, $n \times n$ unknown F but the method trivially extends to rectangular F.

The first term in the two-dimensional problem in (1.14) can be transformed to a one-dimensional problem as before:

$$\|K_2 F K_1^T - \hat{D}\|_{\text{Fro}}^2 = \|Kf - \hat{d}\|^2,$$

where the vectors $f = \text{vec}(F)$ and $\hat{d} = \text{vec}(\hat{D})$ are obtained from matrices F and $\hat{D}$, respectively, and $K = K_1 \otimes K_2$.

Next, let

$$LF = LFI = (I \otimes L)f = L_1 f,$$

$$FL^T = IFL^T = (L \otimes I)f = L_2 f,$$

then

$$\begin{aligned} \|LF\|_{\text{Fro}}^2 + \|FL^T\|_{\text{Fro}}^2 &= \|L_1 f\|^2 + \|L_2 f\|^2 \\ &= f^T L_1^T L_1 f + f^T L_2^T L_2 f \\ &= f^T (L_1^T L_1 + L_2^T L_2) f \\ &= f^T \widetilde{L}^T \widetilde{L} f \\ &= \|\widetilde{L} f\|^2, \end{aligned}$$

where $\widetilde{L}$ is the upper triangular Cholesky factor of the positive definite and symmetric matrix $(L_1^T L_1 + L_2^T L_2)$.

Assume for the moment that $\widetilde{L}^{-1}$ is positive in the sense that $\widetilde{L}^{-1} g \geq 0$ whenever $g \geq 0$. Then with $g = \widetilde{L} f$ the objective function in (1.14) becomes one-dimensional as:

$$\begin{aligned} \min_{f \geq 0} \left(\|Kf - \hat{d}\|^2 + \lambda^2 \|\widetilde{L} f\|^2 \right) &= \min_{\widetilde{L}^{-1} g \geq 0} \left(\|K \widetilde{L}^{-1} g - \hat{d}\|^2 + \lambda^2 \|g\|^2 \right) \\ &\leq \min_{g \geq 0} \left(\|\widetilde{K} g - \hat{d}\|^2 + \lambda^2 \|g\|^2 \right), \end{aligned}$$

where $\widetilde{K} = K \widetilde{L}^{-1}$. In this case we suggest to take $\text{vec}(\hat{F}) = \widetilde{L}^{-1} \hat{g}$ as an estimate of the minimiser in (1.14).

Regarding the assumptions made in the previous paragraph we note that it is a straightforward calculation to show that $\widetilde{L}^T \widetilde{L} = (L_1^T L_1 + L_2^T L_2)$ is a banded, symmetric, positive definite matrix with non-positive off-diagonal elements. In [9] it is shown that such **Stieltjes** matrices have non-negative (elementwise) inverses. While we have not been able to prove the same thing for the Cholesky factor

$\widetilde{L}$, we believe it to be true for the general class of discrete differentiation operators that we have in mind for applications. In particular, all of our numerical examples have exhibited this property. We suggest that the general fact may already be known in the literature and if not, it would make an interesting problem for future investigation. Perhaps a more interesting and important issue is to show that the value of the optimisation problem above posed in terms of $g \geq 0$ is the same as the value of the $f \geq 0$-problem in order justify our use of $\widetilde{L}^{-1}\hat{g}$ as a rigorous estimate for $\mathrm{vec}(\hat{F})$ above.

Finally, as we have pointed out before, methods which involve unfactorised convolutions are computationally unwieldy and it would be helpful to come up with factored versions of the above transformation.

All of these points merit further investigation.

1.5 Duality

Extending the notation of the previous section we define

$$Q(f) = \frac{1}{2}\|Kf - \hat{d}\|^2 + \frac{1}{2}\lambda^2\|Df\|^2 \tag{1.15}$$

and rewrite the optimisation problem

$$\text{Minimize } Q(f), \text{ subject to } f \geq 0. \tag{1.16}$$

Here we are assuming that the regularisation operator D and regularisation parameter λ have been given to us in advance.

Standard duality analysis and the principle of strong duality implies the Karush-Kuhn-Tucker necessary optimality conditions on f and μ (the Lagrangian dual vector):

$$f \geq 0, \;\; \mu \geq 0, \;\; \sum_i \mu_i f_i = 0, \;\; \nabla Q(f) = \mu.$$

A straightforward calculation gives

$$\nabla Q(f) = K^T(Kf - \hat{d}) + \lambda^2 D^T D f.$$

Substituting this result into the KKT conditions yields our basic optimality conditions

$$\begin{aligned} K_i^T(Kf - \hat{d}) = \;\; & -\lambda^2 D_i^T D f \quad \text{if } \; f_i > 0 \\ K_i^T(Kf - \hat{d}) \geq \;\; & -\lambda^2 D_i^T D f \quad \text{if } \; f_i = 0. \end{aligned} \tag{1.17}$$

Here $K_i, \; D_i =$ denote the i-th columns respectively. An important point to be made here is that (1.17) are equivalent to the KKT-optimality conditions.

It is possible to write conditions (1.17) as a closed form expression involving f. First we write

$$g = \lambda^{-2}(\hat{d} - Kf) \tag{1.18}$$

after which we find

$$D^T D f = \max[0, K^T g].$$

Consider now the case of first order regularisation where $D = I$. Then

$$f = \max[0, K^T g]$$

and it is tempting to attempt to recover f via an iterative scheme. However, in practice this approach leads to serious convergence problems as described in [10]. It is exactly at this point that the BRD method [1] provides a way to avoid a direct iterative approach. Details are to be found in [10].

The proposor has asked if an iterative method can be salvaged or, failing that can the BRD method be applied when $D \neq I$. We were not able to answer this question clearly during the week of the workshop, however we record here for completeness some observations made by both the workshop members and the problem proposor.

First, suppose we define $\Gamma = \mathrm{diag}(\|D_i\|^2)$, the diagonal elements of $D^T D$. Γ is strictly positive on the diagonal. Writing $D^T D f = (D^T D - \Gamma) f + \Gamma f$ we can rewrite the above closed form expression as

$$f = \max[0, \Gamma^{-1}(K^T g - (D^T D - \Gamma) f)]. \tag{1.19}$$

Here we are using the fact that the $\max$ operator commutes with Γ.

If we denote by $\tilde{f}$ the least squares best (unregularised) solution we can further simplify (1.19) as

$$f = \max[0, \Gamma^{-1}(\lambda^{-2} K^T K(\tilde{f} - f) - (D^T D - \Gamma) f]$$

so the iterative properties of the map

$$f \to \Gamma^{-1}\lambda^{-2} K^T K \tilde{f} - \Gamma^{-1}(\lambda^{-2} K^T K + D^T D - \Gamma) f$$

need to be explored. In our opinion, the main barrier to convergence is the nonlinear effect invoked by the $\max$ operator in the above iterative scheme.

1.6 Conclusions and Future Work

In this report we have shown how relatively simple, off-the-shelf code can be effectively applied to solve Fredholm Integrals of the first kind through TSVD and Tikhonov regularisation. Higher-order regularisation can also be incorporated with some additional technical difficulties, depending on the nature of the regularising operator. Iterative schemes for solving regularised problems are known in the literature, but work remains in order apply these ideas to the present setting.

Future Work

Bidiagonalisation vs SVD: Eldén's bidiagonalisation algorithm [2] for computing Tikhonov regularised solutions is normally faster than the SVD-based formula (1.8). Developing a version of Eldén's algorithm that exploits the Kronecker product structure would be a good research topic. The work of Faucett and Fulton [4] could be a starting point. However, we expect that a Matlab implementation (without MEX files) of such an algorithm would probably not be any faster than the SVD-based algorithm presented here.

Factored form of higher-order regularisation: It should be straightforward to develop efficient regularisation algorithms when the regularisation operator has factored form. Devising such penalisation operators is an interesting topic for future work. Also, for more general regularisations, the connection between the first order transformed problem and the higher-order problem should be investigated.

Nonnegative constraints: A number of iterative methods are available for regularisation with nonnegative constraints on the solution [11][chapter 9]. It should be straightforward to recode these algorithms to exploit Kronecker product structure. Again, the key to obtaining efficient code is to eliminate expensive Kronecker products from formulae appearing in the algorithm. For example, the gradient projection method involves the objective function and the gradient. The Tikhonov regularisation objective function (1.13) has the gradient

$$K^T(Kf - d) + \lambda^2 f = \mathrm{vec}(K_2^T(K_2 F K_1^T - D)K_1 + \lambda^2 F) \tag{1.20}$$

The right hand side formulae of (1.13) and (1.20) are the ones to use in the iterative algorithm.

Bibliography

[1] J.P. Butler, J.A. Reeds and S.V. Dawson, Estimating solutions of the first kind integral equations with nonnegative constraints and optimal smoothing, SIAM J. Numer. Anal., vol. 18, no. 3, 1981, pp. 381–397.

[2] L. Eldén, Algorithms for the regularization of ill-conditioned least squares problems, BIT, vol. 17, 1977, pp. 134–145.

[3] A.E. English, K.P. Whittall, M.L.G. Joy and R.M. Henkelman, Quantitative two-dimensional time correlation relaxometry, Magn. Reson. Med. vol. 22, pp. 425-434, 1991.

[4] Donald W. Fausett and Charles T. Fulton, Large least squares problems involving Kronecker products, SIAM J. Matrix Anal. Appl. vol. 15, No. 1, pp. 219–227, 1994.

[5] C.W. Groetsch, The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Pitman, Boston, 1984.

[6] Per Christian Hensen, Regularization Tools, a Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems. `pch@imm.dtu.dk, http://www.imm.dtu.dk/~pch`, 2001.

[7] Roger A. Horn and Charles R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.

[8] Julie Kamm and James G. Nagy, Kronecker product and SVD approximations in image restoration, Linear Algebra and Applications vol. 284, 1998, 177–192.

[9] R. Varga, Matrix iterative analysis, Prentice-Hall, 1962, p. 85.

[10] L. Venkataramanan, Y. Song and M.D. Hürlimann, Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions, IEEE Trans. Sign. Proc., vol. 50, no. 5, 2002, pp. 1017–1026.

[11] Curtis R. Vogel, Computational Methods for Inverse Problems, SIAM, 2002.