Numerical methods and machine learning algorithms for solution of Inverse problems

Larisa Beilina*

Department of Mathematical Sciences, Chalmers University of Technology and Gothenburg University, SE-42196 Gothenburg, Sweden

NFMV020/MMF900

Numerical methods and machine learning algorithms for solution of Inverse problems

Methods for image reconstruction and image deblurring

Lecture 3

What is image deblurring

Image deblurring is recovering the original sharp image using a mathematical model.

Original Image

Blurred Image

Figure: left: exact matrix **X**, right: approximated matrix **B**

The blurring model

- Assume linear blurring operator A which is applied to the exact image $\mathbf{X}$. As a result we will have blurred image $\mathbf{B}$.
- $\mathbf{X}$: $m \times n$ matrix representing the exact image.
- $\mathbf{B}$: $m \times n$ matrix representing the blurred image.
- The forward problem is well-posed: by knowing $\mathbf{X}$ and linear operator A compute the blurred image:

$$A(\mathbf{X}) = \mathbf{B}. \tag{1}$$

- However, the inverse problem: by knowing the blurred image $\mathbf{B}$ compute the original image $\mathbf{X}$, is already an ill-posed problem.

The point-spread function

- Usually, the blurring model is described by a Fredholm integral equation of the first kind which takes the form

$$\int_\Omega K(x, y)\mathbf{X}(x)dx = B(y), y \in \Omega. \tag{2}$$

  Here, Ω is a closed bounded set representing our image domain in $\mathbb{R}^n, n = 2, 3$. It is assumed that the kernel $K(x, y) \in C^k(\Omega \times \Omega)$ is the absolutely integrable known function. Then the equation (2) can be represented in the operator form (1) with the operator A defined as

$$A(\mathbf{X}) := \int_\Omega K(x, y)\mathbf{X}(x)dx. \tag{3}$$

- Information from one part of the image “spills over” to another part
- The process is modeled by a point-spread function (PSF) which is modeled by a Kernel in a linear operator A.
- PSF usually assumed to be space-invariant

Examples of PSF

- *Motion blur*: The point source is smeared into a line. Different choices are available in the literature, see for example,

  M. Tico, M. Vehvilainen, *Estimation of motion blur PSF from differently exposed image frames*, EUSIPCO 2006, Florence, Italy, September 4-8, 2006.

Examples of PSF

- *Motion blur*: The point source is smeared into a line. Different choices are available in the literature, see for example,

  M. Tico, M. Vehvilainen, *Estimation of motion blur PSF from differently exposed image frames*, EUSIPCO 2006, Florence, Italy, September 4-8, 2006.

- *Out-of-focus blur* PSF is given by

$$p_{ij} = \begin{cases} 1/(\pi r^2) & \text{if } (i-k)^2 + (j-l)^2 \leq r^2, \\ 0 & \text{elsewhere,} \end{cases}$$

where (k, l) is the center of PSF, and r is the radius of the blur.

Examples of PSF

- *Motion blur*: The point source is smeared into a line. Different choices are available in the literature, see for example,

  M. Tico, M. Vehvilainen, *Estimation of motion blur PSF from differently exposed image frames*, EUSIPCO 2006, Florence, Italy, September 4-8, 2006.

- *Out-of-focus blur* PSF is given by

$$p_{ij} = \begin{cases} 1/(\pi r^2) & \text{if } (i-k)^2 + (j-l)^2 \leq r^2, \\ 0 & \text{elsewhere,} \end{cases}$$

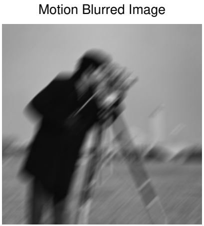
  where (k, l) is the center of PSF, and r is the radius of the blur.

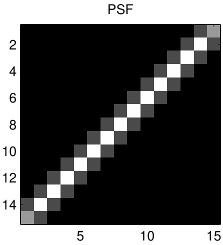
- *Atmospheric turbulence* described as a two-dimensional Gaussian PSF function at every point p_{ij} of PSF:

$$p_{ij} = \exp\left(-\frac{1}{2} \begin{bmatrix} i-k \\ j-l \end{bmatrix}^T \begin{bmatrix} s_1^2 & \rho^2 \\ \rho^2 & s_2^2 \end{bmatrix}^{-1} \begin{bmatrix} i-k \\ j-l \end{bmatrix} \right),$$

  where the parameters s_1, s_2 and ρ determine the width and orientation of the PSF, centered at (k, l).

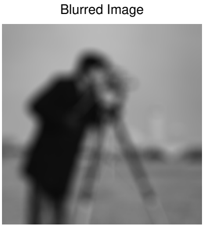
Motion blur

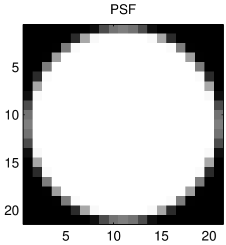




$$A(X) := \int_\Omega K(x,y)X(x)dx = \textit{Blurred} \qquad X$$

Out-of-focus blur



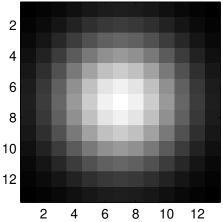


$$A(X) := \int_\Omega K(x,y)X(x)dx = Blurred \qquad X$$

Gaussian blur

Gaussian Blurred Image

PSF



Original Image

$$A(X) := \int_{\Omega} K(x,y)X(x)dx = Blurred \qquad X$$

Image deblurring: solution of an inverse problem

Mathematically, the problem of image deblurring consist in the solution of a Fredholm integral equation of the first kind which is an ill-posed problem. Let H be the Hilbert space H^1 and let $\Omega \subset \mathbb{R}^m, m = 2, 3$, be a convex bounded domain. Our goal is to solve a Fredholm integral equation of the first kind for $x \in \Omega$

$$\int_\Omega K(x-y)z(x)dx = u(y), \tag{4}$$

where $u(y) \in L_2(\bar{\Omega})$, $z(x) \in H$, $K(x-y) \in C^k\left(\overline{\Omega}\right), k \geq 0$ be the kernel of the integral equation.
Let us rewrite (4) in an operator form as

$$A(z) = u \tag{5}$$

with an convolution operator $A : H \to L_2(\bar{\Omega})$ defined as

$$A(z) := \int_\Omega K(x-y)z(x)dx. \tag{6}$$

Image restoration problem

Image restoration problem Let the function $z(x) \in H^1$ of the equation

$$\int_{\Omega} K(x-y)z(x)dx = u(y),$$

be unknown in the domain Ω. Determine the function $z(x)$ for $x \in \Omega$ assuming the functions $K(x-y) \in C^k\left(\overline{\Omega}\right), k \geq 0$ and $u(x) \in L_2(\Omega)$ are known.
The image restoration problem is described by a Fredholm integral equation of the first kind with a linear compact operator $A : H \rightarrow L_2(\bar{\Omega})$. which is an ill-posed problem.
Thus, regularization should be used.

Ill-posed problem: the case of a general compact operator

Let H_1 and H_2 be two Hilbert spaces with $\dim H_1 = \dim H_2 = \infty$. Remind that a sphere in an infinitely dimensional Hilbert space is not a compact set. The Theorem below (see Theorem 1.2 from [BK]) says that the solution of a Fredholm integral equation of the first kind is an ill-posed problem.

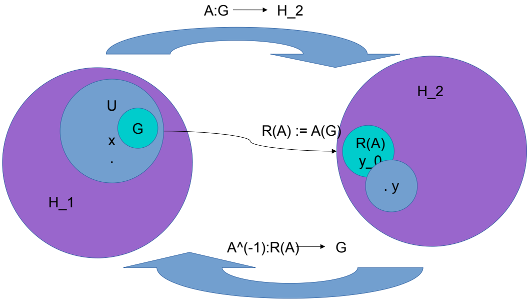
Theorem [BK] *Let* $G = \left\{ \|x\|_{H_1} \leq 1 \right\} \subset H_1$, is not a compact set. *Let* $A : G \to H_2$ *be a compact operator and let* $R(A) := A(G)$ *be its range. Consider an arbitrary point* $y_0 \in R(A)$. *Let* $\varepsilon > 0$ *be a number and* $U_\varepsilon(y_0) = \left\{ y \in H_2 : \|y - y_0\|_{H_2} < \varepsilon \right\}$. *Then there exists a point* $y \in U_\varepsilon(y_0) \diagdown R(A)$. *If, in addition, the operator* A *is one-to-one, then the inverse operator* $A^{-1} : R(A) \to G$ *is not continuous. Hence, the problem of the solution of the equation*

$$A(x) = z, x \in G, z \in R(A) \tag{7}$$

is unstable, i.e. this is an ill-posed problem.

L. Beilina, M. Klibanov, *Approximate global convergence and adaptivity for coefficient inverse problems*, Springer, 2012.

When the problem is an ill-posed problem



Theorem [BK] *Let* $G = \left\{ \|x\|_{H_1} \leq 1 \right\} \subset H_1$, is not a compact set. *Let* $A : G \to H_2$ *be a compact operator and let* $R(A) := A(G)$ *be its range. Consider an arbitrary point* $y_0 \in R(A)$. *Let* $\varepsilon > 0$ *be a number and* $U_\varepsilon(y_0) = \left\{ y \in H_2 : \|y - y_0\|_{H_2} < \varepsilon \right\}$. *Then there exists a point* $y \in U_\varepsilon(y_0) \setminus R(A)$. *If, in addition, the operator* A *is one-to-one, then the inverse operator* $A^{-1} : R(A) \to G$ *is not continuous. Hence, the problem of the solution of the equation* $A(x) = z, x \in G, z \in R(A)$ *is unstable, i.e. this is an ill-posed problem.*

Conclusions about solution of an ill-posed problem

- By the theorem above (see Theorem 1.2 from [BK]) the set $R(A)$ is not dense everywhere. Therefore, the question about the existence of the solution of equation $A(x) = z$ does not make an applied sense. Indeed, since the set $R(A)$ is not dense everywhere, it is very hard to describe a set of values belonging to this set.

- An example, consider the case when the kernel $K(x,y) \in C([a,b] \times [0,1])$ in equation

$$\int_0^1 K(x,y) f(y)\, dy = g(x), x \in (0,1), \tag{8}$$

  is an analytic function of the real variable $x \in (a,b)$. Then the right hand side $g(x)$ of equation (9) should also be analytic with respect to $x \in (a,b)$. However, in applications the function $g(x)$ is a result of measurements, it is given only at a number of discrete points and it definitely contains noise. Clearly it is impossible to determine whether the resulting function is analytic or not.

Conclusions about solution of an ill-posed problem

Conclusion. Assuming that conditions of Theorem above (Theorem 1.2 from [BK]) are satisfied, the problem of solving equation $A(x) = z$ is an ill-posed problem since:

- The proof of an existence theorem makes no applied sense.
- Small fluctuations of the right hand side can lead to large fluctuations of the solution x, i.e. the problem is unstable, see example on the next slide.

Example of instability

Let $\Omega = (0,1)\,, \Omega' = (a,b)\,.$ Let $f_n(x) = f(x) + \sin nx$. Then for $x \in (0,1)$

$$\int_0^1 K(x,y)\, f_n(y)\, dy = \int_0^1 K(x,y)\, f(y)\, dy + \int_0^1 K(x,y) \sin nydy = g_n(x)\,, \tag{9}$$

where $g_n(x) = p(x) + p_n(x)$ and

$$p_n(x) = \int_0^1 K(x,y) \sin nydy.$$

By the Lebesque lemma

$$\lim_{n\to\infty} \|p_n\|_{C[a,b]} = 0.$$

However, it is clear that

$$\left\| f_n(x) - f(x) \right\|_{C[0,1]} = \|\sin nx\|_{C[0,1]}$$

is not small for large n.

Tikhonov functional

Let $\delta > 0$ be the error in the right-hand side of the equation (4):

$$A\left(z^*\right) = u^*, \quad \|u - u^*\|_{L_2(\sigma)} \leq \delta. \tag{10}$$

where u^* is the exact right-hand side corresponding to the exact solution z^*.
To find the approximate solution of the equation (4) corresponding to applications in image restoration usually is minimized the functional

$$M_\alpha\left(z\right) = \|Az - u\|^2_{L_2(\Omega)} + \alpha\|z\|^2_{H_1(\Omega)}, \tag{11}$$

$$M_\alpha : H^1 \to \mathbb{R},$$

where $\alpha = \alpha\left(\delta\right) > 0$ is the small regularization parameter.

Convolution equation

Let us consider an important class of Fredholm integral equations of the first kind - the convolution equation. These equations can be presented in the form (5) with the convolution operator $A : H^1 \to L_2(\Omega)$ defined by

$$A(z) := \int_{\Omega} \rho(y-x) z(x) dx, \tag{12}$$

where $\rho(y-x) \in C^k(\Omega \times \Omega), k \geq 0, \;\; z(x) \in H^1$. Then using the convolution theorem and properties of the Fourier transform we obtain the minimum $z(x) \in H^1$ of the functional (11) given by

$$z(x) = F^{-1}\Big(\frac{\hat{u}(\omega)\hat{\rho}^*(\omega)}{|\hat{\rho}(\omega)|^2 + \alpha(1+\omega^2)}\Big), \tag{13}$$

where $\hat{f}(\omega)$ denotes the Fourier transform $F(f)(\omega)$ of the function $f(\omega)$ defined by

$$\hat{f}(\omega) := F(f)(\omega) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(x) e^{-i\omega x} dx. \tag{14}$$

We consider now more general form of the Tikhonov functional (11). Let W_1, W_2, Q be three Hilbert spaces, $Q \subseteq W_1$ as a set, the norm in Q is stronger than the norm in W_1 and $\overline{Q} = W_1$, where the closure is understood in the norm of W_1. We denote scalar products and norms in these spaces as

$$(\cdot,\cdot), \|\cdot\| \text{ for } W_1,$$

$$(\cdot,\cdot)_2, \|\cdot\|_2 \text{ for } W_2$$

$$\text{and } [\cdot,\cdot], [\cdot] \text{ for } Q.$$

Let $A : W_1 \to W_2$ be a bounded linear operator. Our goal is to find the function $z(x) \in Q$ which minimizes the Tikhonov functional

$$E_\alpha(z) : Q \to \mathbb{R}, \tag{15}$$

$$E_\alpha(z) = \frac{1}{2}\|Az - u\|_2^2 + \frac{\alpha}{2}[z - z_0]^2, u \in W_2; z, z_0 \in Q, \tag{16}$$

where $\alpha \in (0,1)$ is the regularization parameter. To do that we search for a stationary point of the above functional with respect to z satisfying $\forall b \in Q$

$$E'_\alpha(z)(b) = 0. \tag{17}$$

The following lemma is well known [BKK] for the case $W_1 = W_2 = L_2$.
Lemma 1. *Let* $A : L_2 \to L_2$ *be a bounded linear operator. Then the Fréchet derivative of the functional (11) is*

$$E'_\alpha(z)(b) = (A^*Az - A^*u, b) + \alpha[z - z_0, b], \forall b \in Q. \tag{18}$$

In particular, for the integral operator (4) we have

$$\begin{aligned} E'_\alpha(z)(b) = \int\limits_\Omega b(s)\Bigg[\int\limits_\Omega z(y)\Bigg(\int\limits_\Omega K(x-y)K(x-s)dx\Bigg)dy \\ - \int\limits_\Omega K(x-s)u(x)\,dx\Bigg]ds \\ + \alpha[z - z_0, b], \forall b \in Q. \end{aligned} \tag{19}$$

[BKK] A. B. Bakushinsky, M. Y. Kokurin, A. Smirnova, *Iterative methods for ill-posed problems*, Walter de Gruyter GmbH&Co., 2011.

Lemma 2 is also well known, since $A : W_1 \to W_2$ is a bounded linear operator. We formulate this lemma only for our specific case.

Lemma 2. *Let the operator $A : W_1 \to W_2$ satisfies conditions of Lemma 1. Then the functional $E_\alpha(z)$ is strongly convex on the space Q with the convexity parameter κ* such that

$$\left(E'_\alpha(x) - E'_\alpha(z), x - z\right) \geq \kappa[x - z]^2, \forall x, z \in Q. \tag{20}$$

It is known from the theory of convex optimization that Lemma 2 implies existence and uniqueness of the global minimizer $z_\alpha \in Q$ of the functional J_α defined in (15) such that

$$J_\alpha(z_\alpha) = \inf_{z \in Q} J_\alpha(z).$$

It is well known that the operator $F = Az - u$ is Lipschitz continuous

$$\|F(z_1) - F(z_2)\| \leq \|A\| \cdot \|z_1 - z_2\| \ \ \forall z_1, z_2 \in H.$$

We also introduce new constant $D = D\left(\|A\|, \alpha\right) = const. > 0$ such that

$$\left\|J'_\alpha\left(z_1\right) - J'_\alpha\left(z_2\right)\right\| \leq D\left\|z_1 - z_2\right\|, \forall z_1, z_2 \in H. \tag{21}$$

Similarly, the functional $M_\alpha(z)$ is also strongly convex on the Sobolev space H_1:

$$\left(M'_\alpha(x) - M'_\alpha(z), x - z\right)_{H_1} \geq \kappa \|x - z\|^2_{H_1}, \forall x, z \in H_1, \tag{22}$$

To find minimum of (16) or (11) (difference in reg.terms), we can use any gradient-like method. For example, perform usual gradient update to find minimum of (16):

$$z^{k+1} = z^k + \beta E'_\alpha\left(z^k\right)(b).$$

or to find minimum of (11):

$$z^{k+1} = z^k + \beta M'_\alpha\left(z^k\right)(b).$$

until $\|z^{k+1} - z^k\|$ converges.

Bayesian approach

- Assume that measured data y is random because of noise, also assume that the reconstruction x is random. Then it can be used posterior distribution denoted by $P(\mathbf{x} = x | \mathbf{y} = y) := P(x|y)$. Posterior distribution $P(x|y)$ is the probability (density) if the random variable $\mathbf{x}$ is equal to x given that a random variable $\mathbf{y}$ is equal to y.
- Bayes theorem

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}.$$

- Here, $P(y|x)$ is data likelihood - the probability of a measured data y given a reconstruction x, it is defined by the physical model problem. Example: for the Radon transform with Poisson noise the data likelihood is approximated using normal distribution

$$P(y|x) \propto e^{-\frac{1}{2}||Ax-y||^2_{\Lambda^{-1/2}}}$$

with a covariance approximated as $\Lambda = (y_0 e^{-Ax})$, y_0 is the number of photons to heat every pixel in homo media.

Bayesian approach

- $P(x)$ is the prior model, or probability of the how the reconstruction should look like. It should include all information which we know about the reconstruction.
- As soon as $P(y|x)$ and $P(x)$ are defined, the estimator of data likelihood $P(x|y)$ should be chosen. The most popular one is the Maximum a posterior estimate $MAP(y)$ which is the solution of the problem
$$MAP(y) = \max_{x \in X} P(x|y).$$
Taking the log of the above equation, we get
$$\begin{aligned} \max_{x \in X} \log P(x|y) &= \max_{x \in X} [\log P(y|x) + \log P(x) - \log P(y)] \\ &= \min_{x \in X} -[\log P(y|x) + \log P(x)] \end{aligned} \tag{23}$$

Bayesian approach: an example

Let us consider an example of a Gaussian prior $P(x)$ with mean μ and covariance Σ. Then the prior likelihood is:

$$
\begin{aligned}
P(x) &\propto e^{-\frac{1}{2}\|x-\mu\|^2_{\Sigma^{-1/2}}}, \\
\ln P(x) &= -\frac{1}{2}\|x-\mu\|^2_{\Sigma^{-1/2}} + Const.
\end{aligned}
\tag{24}
$$

If we assume the Gaussian noise with mean zero, then the data has mean $A(x)$ and covariance Λ, and data likelihood $P(y|x)$ can be modeled as

$$
\begin{aligned}
P(y|x) &\propto e^{-\frac{1}{2}\|Ax-y\|^2_{\Lambda^{-1/2}}}, \\
\ln P(y|x) &= -\frac{1}{2}\|Ax-y\|^2_{\Lambda^{-1/2}} + Const.
\end{aligned}
\tag{25}
$$

Combining models for $P(x)$ (24) and for $P(y|x)$ (25) into the maximum a posteriori estimate $MAP(y)$ (30), we get

$$
\begin{aligned}
\max_{x \in X} \ln P(x|y) &= \min_{x \in X} -[\ln P(y|x) + \ln P(x)] \\
&= \min_{x \in X} [\frac{1}{2}\|Ax-y\|^2_{\Lambda^{-1/2}} + \frac{1}{2}\|x-\mu\|^2_{\Sigma^{-1/2}}].
\end{aligned}
\tag{26}
$$

Bayesian approach: an example

We can approximate the operator A with a matrix $\mathcal{A}$, take the gradient of (26) we get:

$$
\begin{aligned}
(\log P(x|y))'_x &= (\frac{1}{2}\|\mathcal{A}x - y\|^2_{\Lambda^{-1/2}} + \frac{1}{2}\|x - \mu\|^2_{\Sigma^{-1/2}})'_x \\
&= \mathcal{A}^T\Lambda^{-1}(\mathcal{A}x - y) + \Sigma^{-1}(x - \mu) \\
&= \mathcal{A}^T\Lambda^{-1}\mathcal{A}x - \mathcal{A}^T\Lambda^{-1}y + \Sigma^{-1}x - \Sigma^{-1}\mu \\
&= (\mathcal{A}^T\Lambda^{-1}\mathcal{A} + \Sigma^{-1})x - (\mathcal{A}^T\Lambda^{-1}y + \Sigma^{-1}\mu).
\end{aligned}
\tag{27}
$$

Thus, we obtain system of linear equations considering $(\log P(x|y))'_x = 0$:

$$
(\mathcal{A}^T\Lambda^{-1}\mathcal{A} + \Sigma^{-1})x = \mathcal{A}^T\Lambda^{-1}y + \Sigma^{-1}\mu \tag{28}
$$

The system above can be solved as

$$
x = ((\mathcal{A}^T\Lambda^{-1}\mathcal{A} + \Sigma^{-1}))^{-1}(\mathcal{A}^T\Lambda^{-1}y + \Sigma^{-1}\mu). \tag{29}
$$

Bayesian approach: an example of Spike and Slab model

- Another type of prior which can be taken into consideration, is sparsity inducing prior. This prior is considered when the obtained solution is very smooth.
- We assume that we have discrete values of indicator function $\gamma_j, j = 1, ..., N$ and we consider $\gamma_j = 1$ when the feature j is in our model, and $\gamma_j = 0$ when the feature j is out our model.
- Now let us consider the way to induce sparsity in a specific model parameter and use “Spike and Slab” prior distribution: it is a mixture of a point mass at 0 (to exclude feature), and flat prior (usually Gaussian) for included variables.
- In a mixture model, we consider the Bernoulli random vector $\gamma \in R^p$: when $\gamma_j = 0$ then feature coefficients $\beta_j = 0$, when $\gamma_1 = 1$ then corresponding β_j are taken from Gaussian distribution with a large variance term.

Bayesian approach: an example of Spike and Slab model

Then

$$p(\gamma|\pi_0) = \prod_{j=1}^{p} Bern(\gamma_j|\pi_0) = \pi_0^{||\gamma||_0}(1-\pi_0)^{p-||\gamma||_0}. \tag{30}$$

Here, $||\cdot||_0$ is l_0 norm or number of non-zero elements in the vector, $Bern(\gamma_j|\pi_0)$ are Bernoulli distributions of γ_j with probability π_0.

$$\begin{aligned}\log p(\gamma|\pi_0) &= ||\gamma||_0 \log \pi_0 + (p - ||\gamma||_0)\log(1-\pi_0) \\ &= ||\gamma||_0 \underbrace{(\log \pi_0 - \log(1-\pi_0))}_{\lambda} + Const. \\ &= \lambda||\gamma||_0 + Const.\end{aligned} \tag{31}$$

where $Const. = p\log(1-\pi_0)$. Here, λ controls how sparse the features are. Here, $||\cdot||_0$ is l_0 norm or number of non-zero elements in the vector. Here, π_0 controls how sparse are the features: when $\pi_0 \approx 0$ then $\lambda < 0$ and the feature vector is very sparse.

Bayesian approach: an example of sparcity inducing prior

Using (31) we can write the sparsity inducing prior corresponding to the prior likelihood in MAP (30)

$$\max_{x \in X} \log P(x|y) = \max_{x \in X}[\log P(y|x) + \log P(x) - \log P(y)]$$
$$= \min_{x \in X} -[\log P(y|x) + \log P(x)]$$

is:

$$-\log P(x) = \lambda ||Dx||_0 + Const., \tag{32}$$

where D is an operator such that Dx is sparse, see Spike and Slab model above.

Further reading about Bayesian approach in inverse problems:
J. Kaipio, E. Somersalo, Statistical and computational inverse problems,*Springer*, 160, 2006.

A. M. Stuart, Inverse Problems: a Bayesian perspective, *Acta Numerica*, 19, 451-559, 2010.

The finite element method

- We discretize the bounded domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$ by an mesh T using non-overlapping elements K. $\mathbb{R}^3$ are tetrahedrons, in $\mathbb{R}^2$ the elements K are triangles such that $T = K_1, ..., K_l$, where l is the total number of elements in Ω.
- Let the mesh function $h = h(x)$ is a piecewise-constant function such that
$$h(x) = h_K \quad \forall K \in T,$$
where h_K is the diameter of K which we define as the longest side of K.
- We introduce now the finite element space V_h as
$$V_h = \left\{ v(x) \in V : v \in C(\Omega),\ v|_K \in P_1(K)\ \forall K \in T \right\}, \tag{33}$$
where $P_1(K)$ denotes the set of piecewise-linear functions on K with
$$V = \left\{ v(x) : v(x) \in H^1(\Omega) \right\}.$$
The finite dimensional finite element space V_h is constructed such that $V_h \subset V$.

The finite element method

Let us consider the Tikhonov functional

$$M_\alpha(z) = \frac{1}{2} \|Az - u\|^2_{L_2(\Omega_\kappa)} + \frac{\alpha}{2} |||\nabla z|||^2_{L^2(\Omega)} \, . \qquad (34)$$

- For (34) the CG(1) finite element method reads: find $z_h \in V_h$ such that
$$M'_\alpha(z_h)(b) = 0 \ \ \forall b \in V_h. \qquad (35)$$
- Then the Fréchet derivative of the functional (34) is derived in Lemma 2 of [BGN]

  Lemma 2. Let $A : H^1(\Omega) \to L_2(\Omega_\kappa)$ be a bounded linear operator. Then the Fréchet derivative of the functional (34) is

$$M'_\alpha(z)(b) = (A^*Az - A^*u, b) + \alpha(|\nabla z|, |\nabla b|), \ \forall b \in H^1(\Omega), \qquad (36)$$

  with a convex growth factor b, i.e., $|\nabla b| < b$.

[BGN] Beilina L., Guillot G., Niinimäki K. The Finite Element Method and Balancing Principle for Magnetic Resonance Imaging. In *Mathematical and Numerical Approaches for Multi-Wave Inverse Problems. CIRM 2019.* Springer Proceedings in Mathematics Statistics, vol 328. Springer, Cham (2020) https://doi.org/10.1007/978-3-030-48634-1_9

A posteriori ananlysis in a finite element method

- Let $P_h : V \to M$ for $\forall M \subset V$, be the operator of the orthogonal projection of V on M. Let the function $f \in H^1(\Omega) \cap C(\Omega)$ and $\partial_{x_i} f_{x_i} \in L_\infty(\Omega)$. We define by f_k^I the standard interpolant [EEJ] on triangles/tetrahedra of the function $f \in H$.
- Then by one of properties of the orthogonal projection

$$\|f - P_h f\|_{L_2(\Omega)} \leq \left\|f - f_k^I\right\|_{L_2(\Omega)}. \tag{37}$$

- It follows from formula 76.3 of [EEJ] that

$$\|f - P_h f\|_{L_2(\Omega)} \leq C_I \|h \nabla f\|_{L_2(\Omega)}, \forall f \in V. \tag{38}$$

where $C_I = C_I(\Omega)$ is positive constant depending only on the domain Ω.

[EEJ] K. Eriksson, D. Estep and C. Johnson, *Calculus in Several Dimensions*, Springer, Berlin, 2004.

A posteriori analysis in a finite element method

Let us present a general framework for a posteriori error estimate for:

- For the error $|M_\alpha(z_\alpha) - M_\alpha(z_h)|$ in the Tikhonov functional (34).
- For the error $|z_\alpha - z_h|$ in the regularized solution z_α of this functional.

Note that

$$M_\alpha(z_\alpha) - M_\alpha(z_h) = M'_\alpha(z_h)(z_\alpha - z_h) + R(z_\alpha, z_h), \tag{39}$$

where $R(z_\alpha, z_h)$ is the second order remainder term. We assume that z_h is located in the small neighborhood of the regularized solution z_α. Thus, the term $R(z_\alpha, z_h)$ is small and we can neglect it.
We now use the Galerkin orthogonality principle

$$M'_\alpha(z_h)(b) = 0 \ \ \forall b \in V_h \tag{40}$$

together with the splitting

$$z_\alpha - z_h = (z_\alpha - z_\alpha^I) + (z_\alpha^I - z_h) \tag{41}$$

where $z_\alpha^I \in V_h$ is the interpolant of z_α, and get the following error representation:

$$M_\alpha(z_\alpha) - M_\alpha(z_h) \approx M'_\alpha(z_h)(z_\alpha - z_\alpha^I). \tag{42}$$

A posteriori error in the regularized solution

Theorem 1 [KB]
Let $z_h \in V_h$ be a finite element approximation of the regularized solution $z_\alpha \in H^2(\Omega)$ on the finite element mesh T with the mesh function h. Then there exists a constant D defined by (21) such that the following a posteriori error estimate for the regularized solution z_α holds

$$||z_h - z_\alpha||_{H^1(\Omega)} \leq \frac{D}{\alpha} C_I ||h z_h||_{L_2(\Omega)}.$$

[KB] N. Koshev and L. Beilina, An adaptive finite element method for Fredholm integral equations of the first kind and its verification on experimental data, in the Topical Issue "Numerical Methods for Large Scale Scientific Computing" of CEJM, 11(8), 1489-1509, 2013.

A posteriori error in the Tikhonov functional

Theorem 2 [KB]

Suppose that there exists minimizer $z_\alpha \in H^2(\Omega)$ of the functional M_α on the set V and mesh T. Suppose also that there exists finite element approximation z_h of a minimizer z_α of M_α on the set V_h and mesh T with the mesh function h. Then the following approximate a posteriori error estimate for the error $e = |M_\alpha(z_\alpha) - M_\alpha(z_h)|$ in the Tikhonov functional holds

$$e = |M_\alpha(z_\alpha) - M_\alpha(z_h)| \leq C_I \left\| M'_\alpha(z_h) \right\|_{H^1(\Omega)} \|hz_h\|_{L_2(\Omega)}. \qquad (43)$$

[KB] N. Koshev and L. Beilina, An adaptive finite element method for Fredholm integral equations of the first kind and its verification on experimental data, in the Topical Issue “Numerical Methods for Large Scale Scientific Computing” of CEJM, 11(8), 1489-1509, 2013.

AFEM algorithm

Step 0. Choose an initial mesh T_0 in Ω and obtain the numerical solution z_0 on T_0 using the convolution theorem. Compute the sequence $z_k, k > 0$, on a refined meshes T_k via following steps:

Step 1. Obtain the FEM solution $z_k := z_{hk}$ using the finite element method

$$M'_\alpha(z_h)(b) = 0 \ \ \forall b \in V_h.$$

Step 2. Refine the mesh T_k using the Theorem 1 at all points where

$$\left|z_k\left(x\right)\right| \geq \widetilde{\varkappa}_k \max_\Omega \left|z_k\left(x\right)\right|, \tag{44}$$

or using the Theorem 2 at all points where

$$|M_h(z_k)| \geq \beta_k \max_\Omega |M_h(z_k)|. \tag{45}$$

Here, the tolerances $\widetilde{\varkappa}_k, \beta_k \in (0,1)$ are chosen by the user.

Step 3. Construct a new mesh T_{k+1} in Ω and perform steps 1-3 on the new mesh. Stop mesh refinements when $||z_k - z_{k-1}|| < \epsilon$ or $||M_h(z_k)|| < \epsilon$, or image deteriortaed.

Image deblurring: example

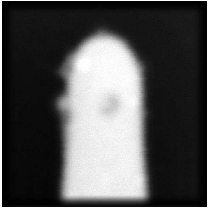
a)

b)

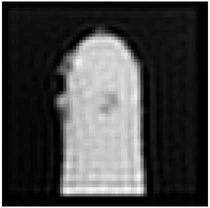
Figure: a) Image of the defect in the planar chip. b) result of reconstruction with bounded total variation functions. Source: [1].

[1] Koshev N.A., Orlikovsky N.A., Rau E.I., Yagola A.G. Solution of the inverse problem of restoring the signals from an electronic microscope in the backscattered electron mode on the class of bounded variation functions, Numerical Methods and Programming, 2011, V.11, pp. 362-367 (in Russian).

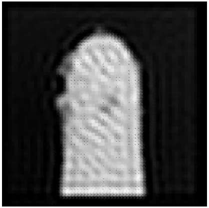
Image deblurring: an adaptive refinement, example



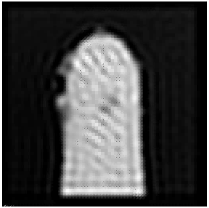
a) 7938 elements



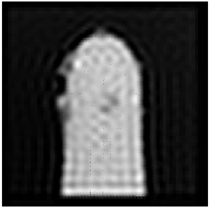
b) z_1, 11270 elements



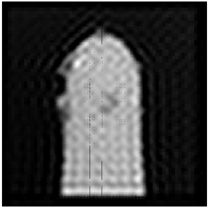
c) z_2, 15916 elements



d) z_3, 24262 elements



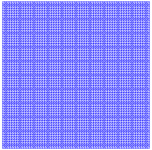
e) z_4, 40358 elements



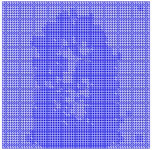
f) z_5, 72292 elements

Figure: Reconstruction from the experimental backscattering data obtained by the microtomograph [KB].

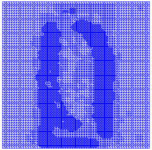
Image deblurring: an adaptive refinement, example



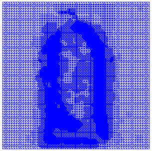
g) 7938 elements



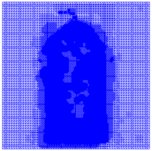
h) 11270 elements



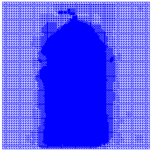
i) 15916 elements



j) 24262 elements



k) 40358 elements



l) 72292 elements

Figure: Reconstruction from the experimental backscattering data obtained by the microtomograph [KB].

[KB] N. Koshev and L. Beilina, An adaptive finite element method for Fredholm integral equations of the first kind and its verification on experimental data, in the Topical Issue “Numerical Methods for Large Scale Scientific Computing” of CEJM, 11(8), 1489-1509, 2013.