Parameter Identification in Partial Differential Equations

Martin Burger

http://www.indmath.uni-linz.ac.at/people/burger

Johannes Kepler Universität Linz

Differentiation of data

Not strictly a parameter identification problem, but good motivation.
Appears often as a subproblem.

Given noisy observation of a function *f*,

$$f^\delta(x) = f(x) + n^\delta(x), \qquad x \in [0,1]$$

Find its derivative

Differentiation of data

Pointwise noise at measurement points identically normally distributed with mean 0 and variance δ.

Law of large numbers yields

$$\int_0^1 |n^\delta(x)|^2 \; dx \approx \delta^2$$

Differentiation of data

Example: $n^\delta(x) = \sqrt{2}\delta \sin(2\pi k x)$

$$\frac{df^\delta}{dx}(x) = \frac{df}{dx}(x) + \sqrt{2}2\pi\delta k \sin(2k\pi x)$$

Noise satisfies $\int_0^1 |n^\delta(x)|^2 \; dx = \delta^2$

But error in derivative is large

Differentiation of data

Error:

$$\left(\int_0^1 \left(\frac{df^\delta}{dx}(x) - \frac{df}{dx}(x)\right)^2 \, dx\right)^{1/2} = 2\pi\delta k$$

$$\sup_{x\in[0,1]} \left|\frac{df^\delta}{dx}(x) - \frac{df}{dx}(x)\right| = \sqrt{2}2\pi\delta k$$

Arbitrarily large since *k* can be arbitrarily large

Conclusion 1

Without regularization and without further information, the error between the exact and noisy solution can be arbitrarily large, even if the noise is arbitrarily small.

Differentiation of data

Additional information:
Solution is smooth (e.g. twice differentiable)

Regularization: e.g. smoothing by elliptic PDE

$$-\alpha\frac{d^2 f_\alpha}{dx^2}(x) + f_\alpha(x) = f^\delta(x), \qquad f_\alpha(0) = f_\alpha(1) = 0$$

Variational principle for elliptic PDEs: minimize

$$H_\alpha(f_\alpha) = \int_0^1 (f_\alpha(x) - f^\delta(x))^2 \, dx + \alpha\int_0^1 \left(\frac{df_\alpha}{dx}(x)\right)^2 \, dx$$

Differentiaton of data

Detailed estimate:

Lecture notes, p. 7

Computerized Tomography

(A nice link between Austria & Norway)

Problem: reconstruct a spatial density *f* in a domain *D* from measurements of X-rays traveling through the domain.

X-rays travel along rays, parametrized by distance *s* from origin and its unit normal vector ω

Computerized Tomography

X-ray beam has intensity *I*

Model: decay of itensity in a small distance Δt proportional to *I, f,* and distance

$$\frac{\Delta I(sw + tw^{\perp})}{\Delta t} = -I(sw + tw^{\perp}) f(sw + tw^{\perp})$$

Limit Δt to zero

$$\frac{dI(sw + tw^{\perp})}{dt} = -I(sw + tw^{\perp}) f(sw + tw^{\perp})$$

Computerized Tomography

Measurement: intensity at emitter and detector

$$I_L(s, w) := I(sw + Lw^{\perp}) \text{ and } I_0(s, w) := I(sw)$$

Parameter identification problem for system of ordinary differential equations (first-order)

Overdetermined for given *f* since initial and final value are known

Computerized Tomography

Integrate ODE

$$\ln I(sw + Lw^{\perp}) - \ln I(sw) = -\int_0^L f(sw + tw^{\perp})\, dt$$

Leads to inversion of the Radon transform

$$\ln I_0(s, w) - \ln I_L(s, w) = \mathcal{R} f(s, w)$$

Computerized Tomography

Radon Transform

$$\mathcal{R}f(s,w) = \int_{\mathbb{R}} f(sw + tw^{\perp})\; dt$$

Exact inversion formula by Johann Radon 1917.
SVD computed by McCormick et. al. in 1960s,
Nobel Prize for Medicine in the 1970

Computerized Tomography

Radial symmetry, $r = \sqrt{s^2+t^2}$

$$\mathcal{R}f(s,w_0) = 2\int_s^{\rho} \frac{rF(r)}{\sqrt{r^2-s^2}}\, dr$$

Computerized Tomography

Radial symmetry, $r = \sqrt{s^2+t^2}$

$$\mathcal{R}f(s,w_0) = 2\int_s^{\rho} \frac{rF(r)}{\sqrt{r^2-s^2}}\, dr$$

Abel Integral equation !

SVs decay with half speed as for differentiation

Groundwater Filtration

Identification of diffusivity of sediments from an observation of the piezometric head (describing the flow)

Knowledge of diffusivity allows to draw conclusions about the structure of the groundwater

Groundwater Filtration

Mathematical model

$$-\text{div}\ (a\nabla u) = f$$

Diffusivity *a*, piezometric head *u* (measured), density of water sources and sinks *f*

+ Appropriate boundary conditions (e.g. *u=0*)

Groundwater Filtration

Instead of exact data, we only know noisy measurement

$$u^\delta(x) = u(x) + n^\delta(x), \qquad x \in \Omega$$

Typically the noise *n(x)* comes from identically normally distributed random variables

From a law of large numbers this gives an estimate

$$\int_0^1 |n^\delta(x)|^2\ dx \approx \delta^2$$

Groundwater Filtration

Consider 1D version of inverse problem, for simplicity with $\frac{du}{dx}(0) = 0$

We can integrate the equation to obtain

$$a(x)\frac{du}{dx}(x) = \int_0^x f(y)\ dy$$

Hence, if the derivative of u does not vanish, a is determined uniquely (Identifiability)

Groundwater Filtration

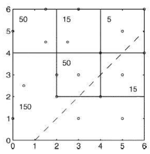
Detailed estimate:

Lecture notes, p. 9

Groundwater Filtration

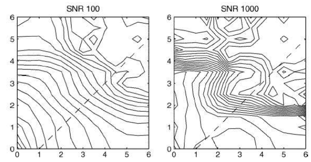
Some results (from Hanke, 1995)

Exact phantom



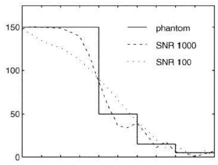
Groundwater Filtration

Reconstructions for two different noise levels (1% and 0.1%)



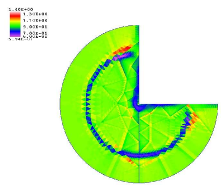
Groundwater Filtration

Reconstruction along the diagonal



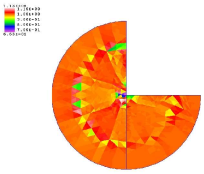
Groundwater Filtration

Reconstructions of piecewise constant parameter, no noise (nonuniqueness curve), fine grid



Groundwater Filtration

Reconstructions of piecewise constant parameter, no noise (nonuniqueness curve), hierarchical grids



Electrical Impedance Tomography

Application in Medical Imaging

Set of electrodes placed around human chest,

Response to various electrical impulses measured

Picture from RPI-EIT Project

Electrical Impedance Tomography

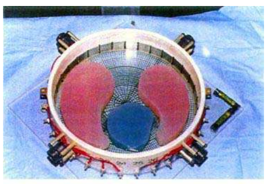
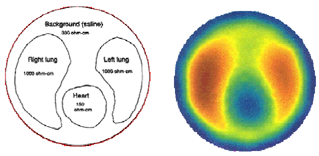


Figure A1. The two-dimensional phantom thorax with pink agar lungs, blue agar heart and black skin in saline. The electrodes are stainless steel, 2.54 x 2.54 cm. The resistivity of the heart is 150 ohm-cm, and that of the lungs is 1000 ohm-cm.

Picture from RPI-EIT Project

Electrical Impedance Tomography



Pictures from RPI-EIT Project

Electrical Impedance Tomography

Mathematical model: Maxwell equations, reduced to potential equation

$$\begin{aligned} \text{div } (a\nabla u) &= 0 \qquad \text{in } D \\ u &= f \qquad \text{on } \partial D \end{aligned}$$

u is electrical potential, *f* applied voltage pattern
a denotes the (unknown) conductivity

Electrical Impedance Tomography

Measurement: current density on the boundary for different voltage patterns

$$g_f = a\frac{\partial u}{\partial n} \qquad \text{on } \partial D$$

Idealized mathematical model: Dirichlet-Neumann map

$$\Lambda_a : f \mapsto g_f$$

Electrical Impedance Tomography

INVERSE CONDUCTIVITY PROBLEM

In practice finite number of measurements

$$\Lambda_a(f_j) \text{ for } j = 1, \ldots, N$$

N typically very large

Electrical Impedance Tomography

To compute output, we have to solve solutions of *N* partial differential equations

$$-\text{ div } (a\nabla u^j) = 0$$

Problem of extremely large scale

Electrical Impedance Tomography

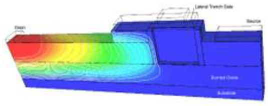
Interesting special case: piecewise constant conductivities

$$a(x) = \begin{cases} a_1 & \text{if } x \in \Omega \subset D \\ a_2 & \text{if } x \in D \backslash \Omega. \end{cases}$$

Real interest is the subset where *a* takes a value different from a_1 (e.g. a tumour)

Semiconductor Devices

Related problem to impedance tomography appears for semiconductor devices: inverse dopant profiling



Inverse Dopant Profiling

Identify the device doping profile from measurements Current-Voltage map

Analogous definitions of voltage and current, but more complicated mathematical model

Mathematical Model

Stationary Drift Diffusion Model:

PDE system for potential *V*, electron density *n* and hole density *p*

$$\begin{aligned} \mathrm{div}(\epsilon_s \nabla V) &= q(n - p - C) \\ \mathrm{div}(D_n \nabla n - \mu_n n \nabla V) &= 0 \\ \mathrm{div}(D_p \nabla p + \mu_p p \nabla V) &= 0 \end{aligned}$$

in Ω (subset of R^2)

Doping Profile *C(x)* enters as source term

Boundary Conditions

Boundary of Ω : homogeneous Neumann boundary conditions on Γ_N and

$$V = V_D = U + V_{bi} = U + U_T \ln\left(\frac{n_D}{n_i}\right)$$

$$n = n_D = \frac{1}{2}\left(C + \sqrt{C^2 + 4n_i^2}\right)$$

$$p = p_D = \frac{1}{2}\left(-C + \sqrt{C^2 + 4n_i^2}\right)$$

on Dirichlet boundary Γ_D (Ohmic Contacts)

Device Characteristics

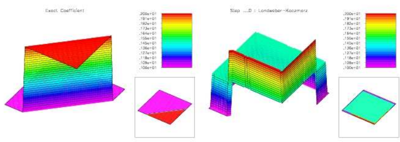
Measured on a contact Γ_0 on Γ_D :

Outflow current density

$$I = (D_n \nabla n - \mu_n n \nabla V - D_p \nabla p - \mu_p p \nabla V) \, .\nu$$

Numerical Tests

Test for a P-N Diode

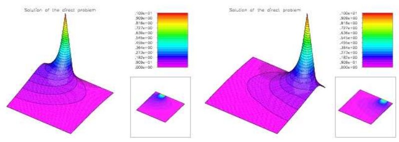


(a) (b)

Real Doping Profile Initial Guess

Numerical Tests

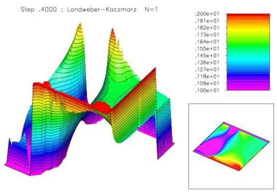
Different Voltage Sources



(a) (b)

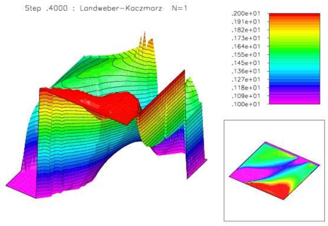
Numerical Tests

Reconstructions with first source

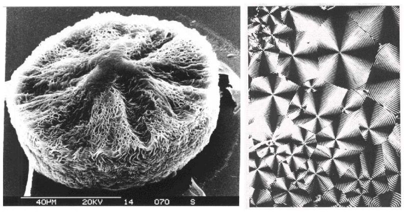


Numerical Tests

Reconstructions with second source



Polymer Crystallization



Polymer Crystallization

Mesoscale model for polymer crystallization

$$\frac{\partial \xi}{\partial t} = \widetilde{G}(T)(1-\xi)u$$
$$\frac{\partial u}{\partial t} = \nabla.(\widetilde{G}(T)v) + \mathcal{F}_d[\widetilde{G}, \widetilde{N}, T]$$
$$\frac{\partial v}{\partial t} = \nabla(\widetilde{G}(T)u)$$
$$c\rho\frac{\partial T}{\partial t} = \nabla.(k\nabla T) + \frac{\partial}{\partial t}(h\xi),$$

ξ = degree of crystallinity

u,v = surface densities

T = temperature

G = growth rate

N = nucleation rate

Polymer Crystallization

Source term

$$\mathcal{F}_1[\widetilde{G},\widetilde{N},T](x,t) := 2\widetilde{N}(T(x,t))_t$$
$$\mathcal{F}_2[\widetilde{G},\widetilde{N},T](x,t) := 2\pi\widetilde{G}(T(x,t))\left(\widetilde{N}(T(x,t)) - \widetilde{N}(T(x,0))\right)$$
$$\mathcal{F}_3[\widetilde{G},\widetilde{N},T](x,t) := 4\pi\widetilde{G}(T(x,t))\int^{t}\widetilde{G}(T(x,s))\left(\widetilde{N}(T(x,s)) - \widetilde{N}(T(x,0))\right)ds$$

Polymer Crystallization

Traditional way of determining nucleation rate as function of temperature:

- Make separate experiment for each value of T
- Count (by eyes !!!) the final number of crystals (typically >> 10^6).
- Divide by volume

Extremely expensive, extremely time-consuming !

Polymer Crystallization

Idea: determine nucleation rate as function of temperature by single nonisothermal experiment

Measured data: temperature T at the boundary of the sample

Degree of crystallinity at final time

Polymer Crystallization

Results

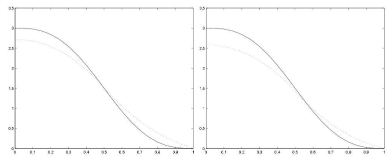


Figure 2: Exact solution (solid) and iterate at $k = k_*$ (dotted) vs. u, $\delta = 0.025$, $f = 50$. The left figure corresponds to an experiment with temperature values in the whole range

Polymer Crystallization

Results

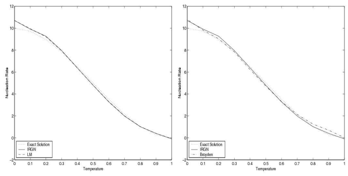


Figure 6: Exact solution (dotted) and closest iterates obtained with the iteratively regularized Gauss-Newton method (solid), the Levenberg-Marquardt method (dashed, left) and Broyden's method (dash-dotted, right), for noise level $\delta = 2\%$.

Polymer Crystallization

Results

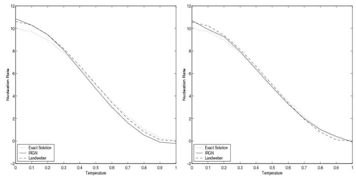
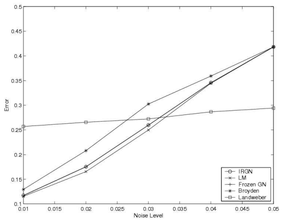


Figure 7: Exact solution (dotted) and closest iterates obtained with the iteratively regularized Gauss-Newton method (solid) and the Landweber iteration (dashed), for noise level $\delta = 5\%$ (left) and $\delta = 2\%$ (right) .

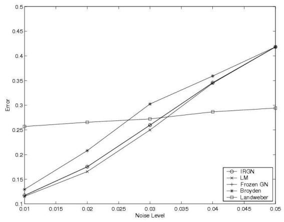
Polymer Crystallization

Reconstruction error vs noise level



Polymer Crystallization

Reconstruction error vs noise level



Polymer Crystallization

Iteration numbers

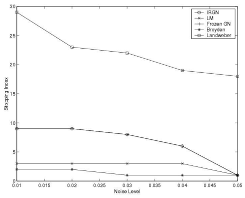


Figure 4: Stopping index $k_*(\delta)$ vs. noise level δ.

Polymer Crystallization

Iteration numbers

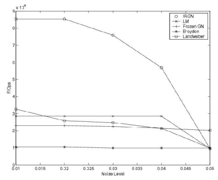


Figure 5: Number of floating point operations $FlOps(k_*(\delta))$ needed until termination of the iteration vs. noise level δ.

Iterative Regularization

For nonlinear inverse problems, iterative regularization is of even higher importance as for linear ones

We need to perform iteration to minimize nonlinear Tikhonov functional anyway

Start from least-squares functional

$$LSQ(x) = \|F(x) - y\|^2$$

Iterative Regularization

Gradient of the least-squares functional

$$LSQ'(x) = 2F'(x)^*(F(x) - y)$$

Perform simple gradient descent:

$$x^{k+1} = x^k - \tau F'(x^k)^*(F(x^k) - y)$$

„Landweber iteration“

Iterative Regularization

Where is the regularization parameter ??!!

Iterative Regularization

Regularization by appropriate early stopping, e.g. discrepancy principle

$$k_* = \inf\{k \in \mathbb{N} \mid \|F(x^k) - y^\delta\| \leq \delta\}$$

Each step of Landweber iteration is well-posed, so we obtain iterative regularization method

Iterative Regularization

Analysis of Landweber iteration: *Hanke-Neubauer-Scherzer* 1994, *Scherzer* 1995

Usual semi-convergence properties as for linear problems, but restriction on the nonlinearity of the problem (replacing regularity of $F'(x^\dagger)$)

Iterative Regularization

Landweber iteration is forward Euler time discretization (with time step τ of the flow)

$$x'(t) = -F'(x(t))^*(F(x(t)) - y)$$

„Asymptotical regularization" (*Tautenhahn* 1994)
Iteration

Iterative Regularization

We can consider other time discretizations:
Backward Euler

$$x^{k+1} = x^k - \tau F'(x^{k+1})^*(F(x^{k+1}) - y)$$

„Iterated Tikhonov Regularization“ (*Hanke-Groetsch* 1999, *Groetsch-Scherzer* 1999)

Iterative Regularization

We can consider other time discretizations:
Semi-Implicit Euler

$$x^{k+1} = x^k - \tau F'(x^k)^*(F(x^k) - y + F'(x^k)(x^{k+1} - x^k))$$

„Levenberg-Marquardt“ (*Hanke* 1995)

Arbitrary Runge-Kutta methods possibly, even inconsistent ones (*Rieder, 2005*)

Iterative Regularization

Levenberg-Marquardt can be rewritten to

$$(F'(x^k)^*F'(x^k) + \alpha^k)(x^{k+1} - x^k) = -F'(x^k)^*(F(x^k) - y)$$

Tikhonov regularization of the linear Newton-equation

$$F'(x^k)(x^{k+1} - x^k) = -(F(x^k) - y)$$

Iterative Regularization

General approach for Newton-type methods:
Apply linear regularization to

$$F'(x^k)(x^{k+1} - x^k) = -(F(x^k) - y)$$

„Iteratively regularized Gauss-Newton“ (*Kaltenbacher-Neubauer-Scherzer* 1996, *Kaltenbacher* 1997)

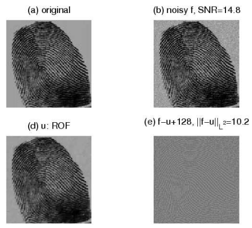
„Newton-CG“ (*Hanke*, 1998)

Iterative Regularization

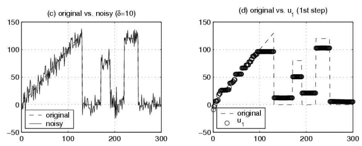
„Newton-Landweber“ (*Kaltenbacher* 1999)
„Broyden“ (*Kaltenbacher*, 1997)
„Multigrid / Discretization“ (*Kaltenbacher*, 2000-2003)

In particular for parameter identification:
„Levenberg-Marquardt SQP“ (*B-Mühlhuber* 2002)
„Newton-Kaczmarcz“ (*B-Kaltenbacher 2004*)

Total Variation Denoising



Total Variation Denoising



TV-Images

Basic problem in **denoising**:
given noisy version g of image f, find „optimal“ approximation u of f

Basic problem in **deblurring**:
given noisy version g of Kf, find „optimal“ approximation u of f

TV-Images

What is optimal ?

Satisfy other criterion, e.g., minimization

$$J(u) \to \min_u$$

over a class of approximations

ROF-Model

What is a suitable class of approximations ?

Discrepancy principle: allow images with

$$\int_\Omega (u-g)^2 \; dx \leq \sigma^2$$

with σ being the noise level:

$$\int_\Omega (f-g)^2 \; dx \leq \sigma^2$$

ROF-Model

Rudin-Osher-Fatemi 1989 (ROF):

Minimize total variation

$$J(u) = \int_\Omega |\nabla u| \; dx$$

subject to

$$\int_\Omega (u-g)^2 \; dx \leq \sigma^2$$

ROF-Model

Deblurring:

Minimize total variation

$$J(u) = \int_\Omega |\nabla u| \; dx$$

subject to

$$\int_\Omega (Ku-g)^2 \; dx \leq \sigma^2$$

ROF-Model

Acar-Vogel 1994:

ROF is regularization method (well-posedness + convergence as σ to zero)

Chambolle-Lions 1997:

Solution unique, there exists Lagrange multiplier λ, problem equivalent to

$$\frac{\lambda}{2}\int_\Omega (u-g)^2\, dx + J(u) \to \min_u$$

ROF-Model

Acar-Vogel 1994:

ROF is regularization method (well-posedness + convergence as σ to zero)

Chambolle-Lions 1997:

Solution unique, there exists Lagrange multiplier λ, problem equivalent to

$$\frac{\lambda}{2}\int_\Omega (u-g)^2\, dx + J(u) \to \min_u$$

Total Variation

Rigorous definition of total variation

$$|u|_{TV} := \sup_{\mathbf{g}\in C_0^\infty(\Omega)^d} \int_\Omega u \text{ div } \mathbf{g}\ dt$$

$$BV(\Omega) := \{\ u \in L^1(\Omega) \mid |u|_{TV} < \infty\ \}.$$

Total Variation

BV includes discontinuous functions:

1D example

$$u^R(x) = \begin{cases} 1 & \text{if } |x| \leq R \\ 0 & \text{else.} \end{cases}$$

Total Variation

1D example

$$\int_\Omega u \text{ div } g \, dt = \int_{-R}^{R} \frac{dg}{dt} \, dt = g(R) - g(-R)$$

$$|u|_{TV} := \sup_{g \in C_0^\infty([-1,1])} [g(R) - g(-R)] = 2$$

Total Variation

Formal optimality for TV-Denoising

$$u - y^\delta = \alpha \text{div} \left(\frac{\nabla u}{|\nabla u|} \right)$$

Term on the right-hand side corresponds to mean curvature of level sets of *u*

Duality

Consider TV-regularization

$$\inf_u \left[\int_\Omega (u - y^\delta)^2 \, dt + \alpha |u|_{TV} \right] =$$

$$\inf_u \sup_{\mathbf{g}} \left[\int_\Omega (u - y^\delta)^2 \, dt + 2\alpha \int_\Omega u \text{ div } \mathbf{g} \, dt. \right]$$

Duality

Exchange inf and sup

$$\sup_{\mathbf{g}} \inf_u \left[\int_\Omega (u - y^\delta)^2 \, dt + 2\alpha \int_\Omega u \text{ div } \mathbf{g} \, dt. \right]$$

Solve inner inf problem for *u*

$$u = y^\delta - \alpha \text{ div } \mathbf{g}$$

Duality

Remains sup problem for ***g***

$$\sup_{\mathbf{g}} \left[\alpha^2 \int_\Omega (\operatorname{div} \mathbf{g})^2 \, dt + 2\alpha \int_\Omega (y^\delta - \alpha \operatorname{div} \mathbf{g}) \operatorname{div} \mathbf{g} \, dt \right]$$

Rewrite equivalently

$$- \int_\Omega (\alpha \operatorname{div} \mathbf{g} - y^\delta)^2 \, dx \to \max_{|\mathbf{g}|_\infty \leq \mathbf{1}}$$

Duality

Introduce $p := \alpha g$

$$\int_\Omega (\operatorname{div} p - y^\delta)^2 \, dt \to \min_{|p|_\infty \leq \alpha}$$

Dual TV problem (*Chambolle 2003*)

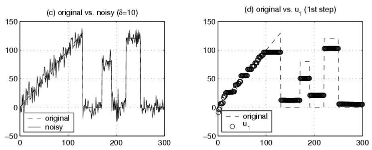
Dual space of BV

$$BV^* := \{ \, q = \operatorname{div} p \mid p \in L^\infty(\Omega) \, \}$$

$$\|q\|_{BV^*} := \inf\{ \|p\|_\infty \mid q = \operatorname{div} p \}$$

Stair-Casing

Lecture notes p.20



ROF-Model

Meyer 2001 (and others before):

ROF has a systematic error, since

$$J(\hat{u}) < J(f)$$

This means that jumps in the reconstructed image are smaller than jumps in the true image

Noise Decomposition

How to resolve the systematic error ?

Take some λ, and minimize TV-functional

$$\frac{\lambda}{2}\int_\Omega (u-g)^2 \, dx + J(u) \rightarrow \min_u$$

to obtain u_1. Then take residual („noise")

$$v_1 = g - u_1$$

Noise Decomposition

Minimize

$$\frac{\lambda}{2}\int_\Omega (u-g-v_1)^2 \, dx + J(u) \rightarrow \min_u$$

to obtain next iterate u_2

The second step may **increase total variation !**

Iterated Total Variation

Obtain new iterate u_k by minimizing

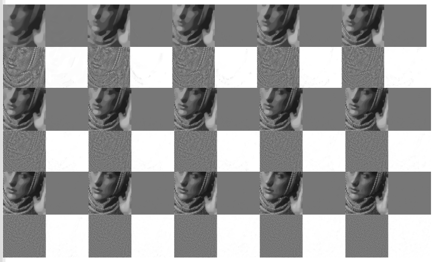
$$\frac{\lambda}{2}\int_\Omega (u-g-v_{k-1})^2 \, dx + J(u) \rightarrow \min_u$$

then decompose noise

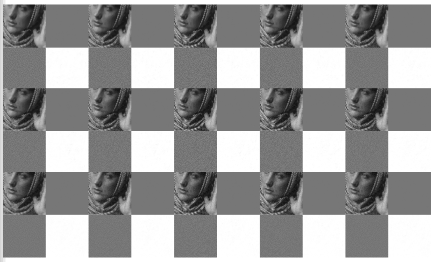
$$v_k = v_{k-1} + f - u_k$$

Need not change the code, just the data !

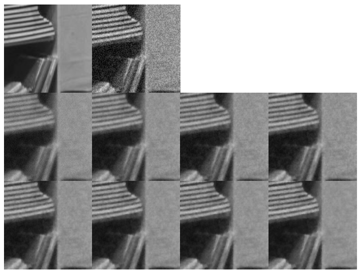
Some Results



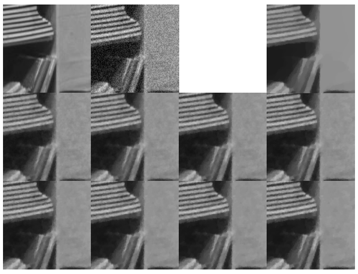
Some Results



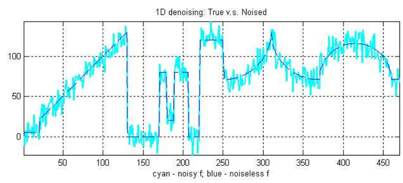
Some Results



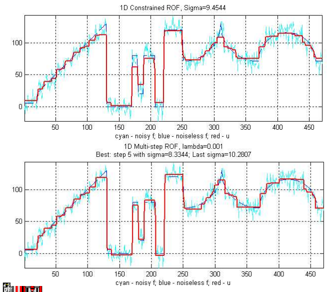
Some Results



Some Results



Some Results



Convergence Analysis

So why does that work ???

Expand fitting term, throw away constant terms:

$$\frac{\lambda}{2}\int_\Omega (u-g)^2 \; dx + J(u) - \int_\Omega v_{k-1}(u-g) \; dx \to \min_u$$

Convergence Analysis

Optimality condition implies

$$v_k = v_{k-1} + f - u_k \in \partial J(u_k)$$

Write equivalent optimization problem

$$\frac{\lambda}{2}\int_\Omega (u-g)^2 \; dx + J(u) - J(u_{k-1}) - \int_\Omega v_{k-1}(u-u_{k-1}) \; dx \to \min_u$$

Convergence Analysis

Hence, iterated TV is „proximal point algorithm"

$$\frac{\lambda}{2}\int_\Omega (u-g)^2 \; dx + D(u, u_{k-1}) \to \min_u$$

Convergence Analysis

The second term is called **Bregman distance**

$$D(u, u_{k-1}) = \\ J(u) - J(u_{k-1}) - \langle v_{k-1}, u - u_{k-1} \rangle$$

with $$v_{k-1} \in \partial J(u_{k-1})$$

Convergence Analysis

Example: quadratic case

$$J(u) = \frac{1}{2} \int_\Omega |\nabla u|^2 \, dx$$

yields

$$D(u, u_{k-1}) = \frac{1}{2} \int_\Omega |\nabla u - \nabla u_{k-1}|^2 \, dx$$

Iterated Tikhonov / Levenberg-Marquardt Method !

Inverse TV Flow

Does the reconstruction depend on λ ?

To understand the dependence on the parameter, consider limit $\lambda \to 0$.

Let $\lambda_N = T/N$, set $t_k = k \lambda_N$. We compute the solution of the iterated TV flow with this specific parameter and set

$$u^N(t_k) = u_k$$

Linear interpolation for other times.

Inverse TV Flow

We obtain a limiting flow,

$$u^N(t) \to u(t)$$

This flow is uniquely defined

Formally, $$\frac{\partial}{\partial t} J'(u(t)) = f - u(t)$$

Inverse TV Flow

The „inverse TV flow“ can be considered as an „inverse scale space method“ (*Scherzer-Weickert 99*)

Start with zero image, and gradually insert smaller and smaller scales. Noise should return in the end, this happened in experiments

Inverse TV Flow

The parameter λ can be interpreted as a time step for an implicit discretization of the inverse TV flow

Should work well and be independent of the parameter, as long as λ is small enough. This can be observed in experiments, too.

Nonlinear Inverse Problems

The algorithm can immediately be generalized to arbitrary linear and even nonlinear inverse problems:

Find next iterate as minimizer of

$$Q_k(u) = \frac{\lambda}{2}||F(u) - g||^2 + D(u, u_{k-1})$$

Nonlinear Inverse Problems

Example: estimate diffusion coefficient *q* in

$$-\text{ div } (q\nabla u) = f$$

from measurement of *u*

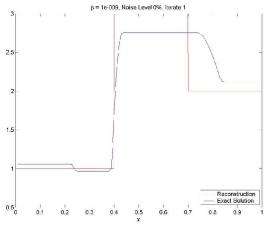
Operator *F: q* $\to$ *u(q)*

Application: find material layers in ground-water modeling (piecewise constant)

Nonlinear Inverse Problems

Exact data

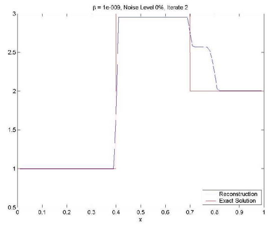
Iterate 1



Nonlinear Inverse Problems

Exact data

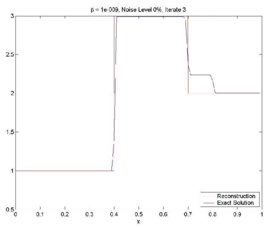
Iterate 2



Nonlinear Inverse Problems

Exact data

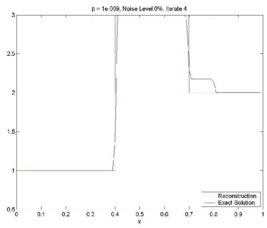
Iterate 3



Nonlinear Inverse Problems

Exact data

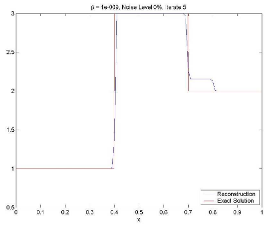
Iterate 4



Nonlinear Inverse Problems

Exact data

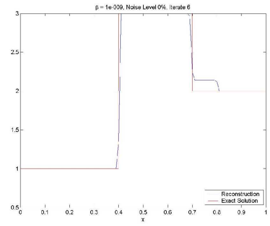
Iterate 5



Nonlinear Inverse Problems

Exact data

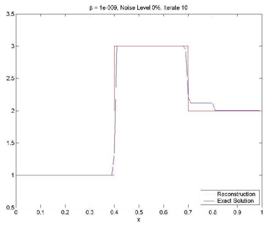
Iterate 6



Nonlinear Inverse Problems

Exact data

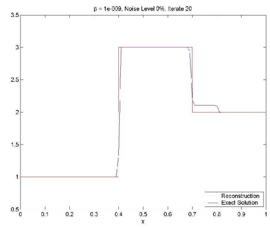
Iterate 10



Nonlinear Inverse Problems

Exact data

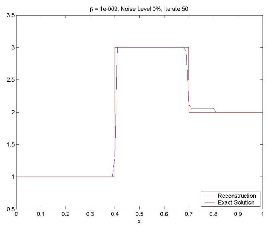
Iterate 20



Nonlinear Inverse Problems

Exact data

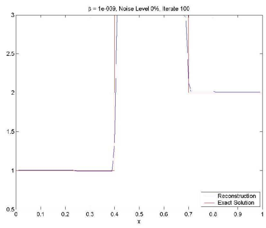
Iterate 50



Nonlinear Inverse Problems

Exact data

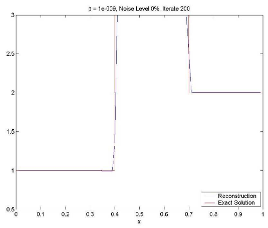
Iterate 100



Nonlinear Inverse Problems

Exact data

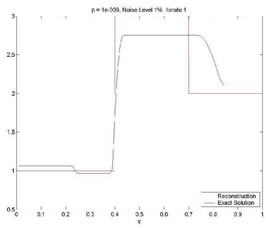
Iterate 200



Nonlinear Inverse Problems

1% noise

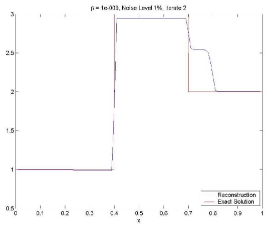
Iterate 1



Nonlinear Inverse Problems

1% noise

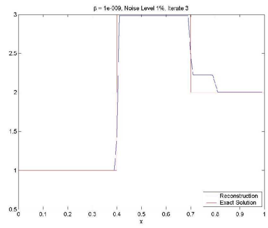
Iterate 2



Nonlinear Inverse Problems

1% noise

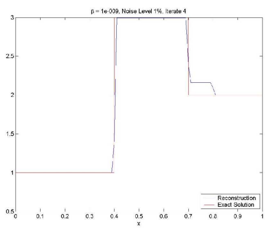
Iterate 3



Nonlinear Inverse Problems

1% noise

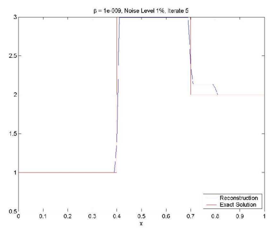
Iterate 4



Nonlinear Inverse Problems

1% noise

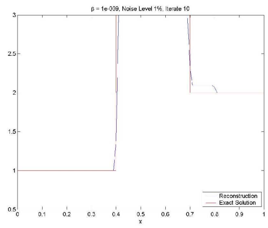
Iterate 5



Nonlinear Inverse Problems

1% noise

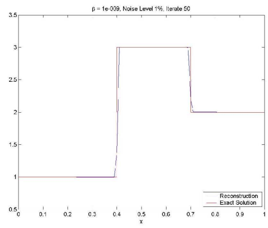
Iterate 10



Nonlinear Inverse Problems

1% noise

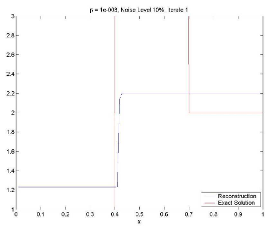
Iterate 50



Nonlinear Inverse Problems

10% noise

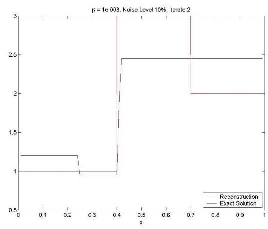
Iterate 1



Nonlinear Inverse Problems

10% noise

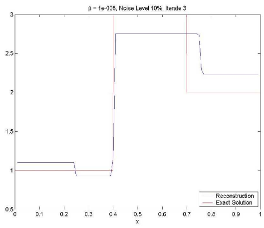
Iterate 2



Nonlinear Inverse Problems

10% noise

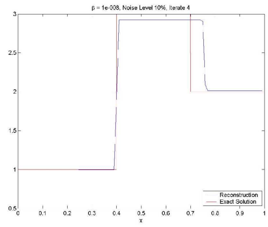
Iterate 3



Nonlinear Inverse Problems

10% noise

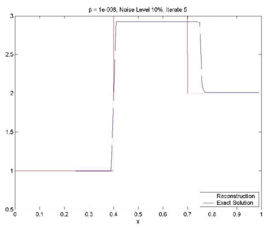
Iterate 4



Nonlinear Inverse Problems

10% noise

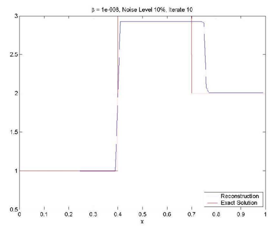
Iterate 5



Nonlinear Inverse Problems

10% noise

Iterate 10



Parameter Identification

Implementation Aspects

Landweber iteration:

- Discretize parameter
- Discretize state + state equation
- Compute gradient by adjoint method
- Try to use same kind of discretization for adjoint as for state equation („Discretized Adjoint = Adjoint of Discretized Problem")
- For multiple state equations compute gradients corresponding to each state equation immediately

Implementation Aspects

Landweber iteration with Black Box Solver:

- Discretize parameter
- Discretize state + state equation (determined by solver)
- If adjoint method not possible, try to compute gradients by finite differencing
- For multiple state equations compute gradients corresponding to each state equation immediately

Implementation Aspects

Quasi-Newton Methods (BFGS):

- Discretize parameter
- Discretize state + state equation (determined by solver)
- Compute gradients by adjoint method, use adjoint to update quasi-Newton matrix
- Solve linear system for update

Implementation Aspects

Newton Methods (LM / IRGN / NCG):

- Discretize parameter
- Discretize state + state equation
- Never compute Newton-Matrix (NOT SPARSE)!!

$$F'(a) = -B \circ \frac{\partial e}{\partial u}(\Phi(a);a)^{-1} \circ \frac{\partial e}{\partial a}(\Phi(a);a)$$

- Try to use iterative method for computing the Newton step, only implement application of $F'(a)$ and its adjoint (two PDEs)

Implementation Aspects

One shot methods:

- Discretize parameter + state + state equation + adjoint + adjoint equation
- Solve the full KKT system simultaneously, e.g.

$$\begin{aligned} 0 &= 2\alpha^k(a^{k+1}-a^k) + \frac{\partial e}{\partial a}(u^k;a^k)^* w^k \\ 0 &= 2B^*(Bu^k - y^\delta + Bv^k) + \frac{\partial e}{\partial a}(u^k;a^k)^* w^k \\ 0 &= \frac{\partial e}{\partial u}(u^k;a^k)v^k + \frac{\partial e}{\partial a}(u^k;a^k)(a^{k+1}-a^k) \end{aligned}$$

Implementation Aspects

One shot methods:

- KKT system is **sparse and symmetric**
- KKT system is not positive definite (no CG !)
- Use BiCGStab, MINRES, QMR or GMRES
- Apply appropriate preconditioning, note that there is a small parameter (α), acting like singular perturbation

Preconditioning strategies: *Battermann-Sachs 2000, Battermann-Heinckenschloss 2001, Ascher-Haber 2001-2003, B-Mühlhuber 2002, Griewank 2005*

Kaczmarz Methods

For problems like impedance tomography, the operators are of the form

$$e(u;a) = (e_1(u^1;a), \ldots, e_N(u^N;a)) = 0$$

$$Bu = (B_1 u^1, \ldots, B_N u^N)$$

Idea: perform multiplicative splitting (like in Gauss-Seidel), cyclic iteration over the *N* subproblems

Kaczmarcz Methods

Subproblem *j*:

$$e_j(u^j; a) = 0, \quad B_j u^j = y_j^\delta$$

Landweber-Kaczmarcz (*Natterer 1996, Kowar-Scherzer 2003*)

$$e(u^{k,j}; a^{k,j-1}) = 0, \qquad \frac{\partial e_j}{\partial u}(u^{k,j}; a^{k,j-1})^* w^{k,j} + B_j^*(B_j u^{k,j} - y_j^\delta) = 0$$

$$a^{k,j} = a^{k,j-1} + \tau^j \frac{\partial e_j}{\partial a}(u^j; a^{k,j-1})^* w^j, \qquad j = 1, \ldots, N$$

Kaczmarcz Methods

Levenberg-Marquardt Kaczmarcz (*B-Kaltenbacher 2004*)

State equation: $e(u^{k,j}; a^{k,j-1}) = 0$

Linear KKT:

$$\begin{aligned}
B^*(Bu^{k,j} + Bv^{k,j} - y^\delta) + \frac{\partial e}{\partial u}(u^{k,j}; a^{k,j-1})^* w^{k,j} &= 0 \\
\alpha_{k,j}(a^{k,j} - a^{k,j-1}) + \frac{\partial e}{\partial a}(u^{k,j}; a^{k,j-1})^* w^{k,j} &= 0 \\
\frac{\partial e}{\partial u}(u^{k,j}; a^{k,j-1}) v^{k,j} + \frac{\partial e}{\partial a}(u^{k,j}; a^{k,j-1})(a^{k,j} - a^{k,j-1}) &= 0
\end{aligned}$$

Kaczmarcz Methods

Simple model problem: identify *q* in

$$-\Delta u_j + q u_j = 0$$

from measurements $f_j = u_j$

for different sources $\dfrac{\partial u_j}{\partial \nu} = g_j$

Numerical Solution

Newton-Kaczmarz approach is equivalent to minimizing

$$\frac{1}{2}\left\|\frac{\partial v_n}{\partial \nu} - g_n\right\|^2_{H^{-1/2}(\partial\Omega)} + \frac{\alpha_n}{2}\|s + q_n - q_0\|^2_{H^1(\Omega)}$$

for $s = q_{n+1} - q_n$ in each step, subject to

$$-\Delta v_n + q_n v_n + s u_n = 0$$

Numerical Solution

Since we need another function to realize the $H^{-1/2}$-norm, corresponding KKT system is 4 x 4

$$-\Delta v_n + q v_n + s u_n = 0$$
$$-\Delta \lambda + q\lambda = 0$$
$$-\Delta \mu + \mu + (1-q) v_n - s u_n = 0$$
$$-\Delta s + s + \frac{1}{\alpha} u_n \lambda = 0$$

Numerical Solution

With boundary conditions

$$v_n = 0$$
$$\lambda - \mu = \phi_n - \phi_{g_n}$$
$$\frac{\partial \mu}{\partial \nu} = 0$$
$$s = 0$$

and

$$\int_\Omega (\nabla \phi_g \cdot \nabla \psi + \phi_g \psi)\ dx = \int_{\partial\Omega} g\psi\ d\sigma \qquad \forall\ \psi \in H^1(\Omega)$$

Numerical Results

Numerical experiment with N=20 localized sources g_j (5 on each side)

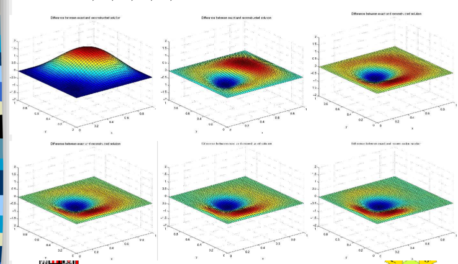
Exact solution

$$\hat{q} = 3 + 5 sin(\pi x) sin(\pi y)$$

Constant initial value $q_0 \equiv 3$

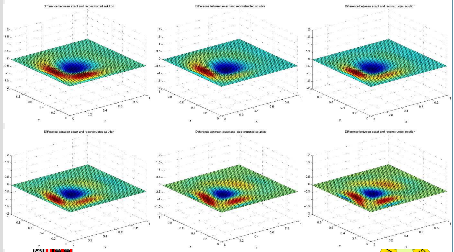
Numerical Results

Iterates 1, 2, 3, 4, 5, 6



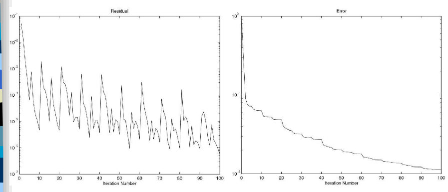
Numerical Results

Iterates 10, 15, 20, 40, 60, 80 (1/2,3/4,1,2,3,4)



Numerical Results

Residuals and Error



Numerical Results

Numerical experiment with N=20 localized sources g_j (5 on each side)

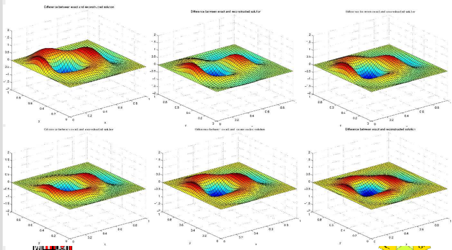
Exact solution

$$\hat{q} = 3 + 5 sin(\pi x) sin(\pi y)$$

Constant initial value $q_0 \equiv 3$

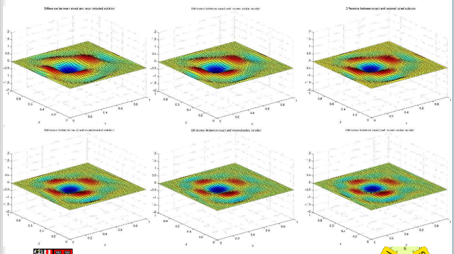
Numerical Results

Iterates 1, 6, 11, 16, 21, 41



Numerical Results

Iterates 80, 120, 160, 240, 320, 400   (4,6,8,12,16,20)



Numerical Results

Residuals and Error