Introduction to Inverse Problems

Rodrigo B. Platte

November 7, 2011
Inverse problems are typically ill-posed. Therefore, further regularization is usually required!
Inverse Problems vs Forward Problems

Question → Answer

Question ← (approximate or incomplete) Answer

Inverse problems are typically ill-posed. Therefore, further regularization is usually required!
Inverse problems are typically ill-posed.
Inverse problems are typically ill-posed. Therefore, further regularization is usually required!
Considering the mapping $A$ which takes the solution $f$ to output data $g$. $Af = g$. Inverse problem: find $f$ given $g$ and $A$.

**Definition (well-posed)**

The problem of finding $f$ from $g$ is called well-posed (Hadamard, 1923) if all

- **Existence** - a solution exists for any data $g$ in data space
- **Uniqueness** - the solution is unique
- **Stability** - continuous dependence of $f$ on $g$: the inverse mapping $g \rightarrow f$ is continuous
Consider the mapping $A$ which takes the solution $f$ to output data $g$. $Af = g$. Inverse problem: find $f$ given $g$ and $A$.

**Definition (well-posed)**

The problem of finding $f$ from $g$ is called well-posed (Hadamard, 1923) if all

- **Existence** - a solution exists for any data $g$ in data space
- **Uniqueness** - the solution is unique
- **Stability** - continuous dependence of $f$ on $g$: the inverse mapping $g \rightarrow f$ is continuous

The first two conditions are equivalent to saying that the operator $A$ has a well defined inverse $A^{-1}$. Moreover, we require that the domain of $A^{-1}$ is all of data space.
Ill-posed problem

**Definition (Ill-Posed: according to Hadamard)**

A problem is ill-posed if it does not satisfy all three conditions for well-posedness. Alternatively an ill-posed problem is one in which

1. \( g \notin \text{range}(A) \)
2. inverse is not unique because more than one image is mapped to the same data, or
3. an arbitrarily small change in the data can cause an arbitrarily large change in the image.
A simple example (discrete)

Consider the linear system

\[
A = \begin{bmatrix}
0.16 & 0.10 \\
0.17 & 0.11 \\
2.02 & 1.29
\end{bmatrix}, \quad b = \begin{bmatrix}
0.26 \\
0.28 \\
3.31
\end{bmatrix}, \quad x = \begin{bmatrix}
1 \\
1
\end{bmatrix}
\]

The least squares solution yields

\[
x_{ls} = [1, 1]^T, \quad \|Ax_{ls} - b\|^2 = 0, \quad \|x - x_{ls}\|^2 = 0
\]

Perturbing \( b \) by \( \delta b = [.01, .01, .001]^T \) gives

\[
x'_{ls} = [1.6857, -0.0718]^T, \quad \|Ax'_{ls} - b\|^2 = 0.0018, \quad \|x - x'_{ls}\|^2 = 1.6189
\]

- A small residual does not imply a realistic solution
- Ill-conditioning of \( A \) leads to a poor solution \( (\kappa(A) = \|A\|\|A^\dagger\|) \)
- Perturbing \( b \) leads to a larger perturbation in \( x \).
Another simple example (scalar)

Let \( f(x) = x \exp(-1/x^2) + 1 \). Find \( x \) such that \( 1 = f(x) \).

Recall definition of (relative) condition number

Forward problem:

\[
\kappa(x) = \frac{\text{Rel Forward Error}}{\text{Rel Backward Error}} = \left| \frac{f(x + \delta x) - f(x)}{f(x)} \right| \approx \left| \frac{xf'(x)}{f(x)} \right|
\]

Inverse problem: \( \kappa(y) \approx \frac{|y|}{\left| (g(y)f'(g(y))) \right|} \) \( (\kappa(1) = \infty) \)
A PDE example (diffusion equation)

\[ u_t = u_{xx}, \quad (t, x) \in (0, \infty) \times (-1, 1) \]

\[ u(t, -1) = u(t, 1) = 0, \quad u(0, x) = u_0(x) \]

Forward problem: given \( u_0 \) find \( u(t, x) \) for some \( t > 0 \).

Backward problem: given \( u(T, x) \) find \( u_0 \).
A PDE example (diffusion equation)

\[ u_t = u_{xx}, \quad (t, x) \in (0, \infty) \times (-1, 1) \]

\[ u(t, -1) = u(t, 1) = 0, \quad u(0, x) = u_0(x) \]

Forward problem: given \( u_0 \) find \( u(t, x) \) for some \( t > 0 \).
Backward problem: given \( u(T, x) \) find \( u_0 \).
Heat equation - Fourier expansion

\[ u_t = u_{xx}, \quad (t, x) \in (0, \infty) \times (-\pi, \pi) \]

\[ u(t, -\pi) = u(t, \pi), \quad u_x(t, -\pi) = u_x(t, \pi) \quad u(0, x) = u_0(x) \]

Let

\[ u(t, x) \approx \sum_{k=-N}^{N} \hat{u}_k(t) \exp(ikx) \]

Then

\[ u_t(t, x) \approx \sum_{k=-N}^{N} \hat{u}_k'(t) \exp(ikx) \quad \text{and} \quad u_{xx}(t, x) \approx \sum_{k=-N}^{N} -k^2 \hat{u}_k(t) \exp(ikx) \]

Therefore, \( \hat{u}_k'(t) = -k^2 \hat{u}_k(t) \) and \( \hat{u}_k(t) = \exp(-k^2 t) \hat{u}_k(0), \)

\[ u(t, x) \approx \sum_{k=-N}^{N} \exp(-k^2 t) \hat{u}_k(0) \exp(ikx) \]
Now suppose that we want to recover $u(0, x)$ from $u(T, x)$. Then $\hat{u}_k(T) = \exp(-k^2T)\hat{u}_k(0)$ gives

$$\hat{u}_k(0) = \exp(k^2T)\hat{u}_k(T).$$

And

$$u(0, x) \approx \sum_{k=-N}^{N} \exp(k^2T)\hat{u}_k(T) \exp(ikx).$$
Now suppose that we want to recover \( u(0, x) \) from \( u(T, x) \).
Then \( \hat{u}_k(T) = \exp(-k^2 T)\hat{u}_k(0) \) gives
\[
\hat{u}_k(0) = \exp(k^2 T)\hat{u}_k(T).
\]

And
\[
u(0, x) \approx \sum_{k=-N}^{N} \exp(k^2 T)\hat{u}_k(T) \exp(ikx).
\]

In the presence of noise,
\[
u_{\text{recovered}}(0, x) = \sum_{k=-N}^{N} \exp(k^2 T)(\hat{u}_k(T) + \eta_k)) \exp(ikx)
\]
\[
= u(0, x) + \sum_{k=-N}^{N} \exp(k^2 T)\eta_k \exp(ikx)
\]

Noise in highest modes are amplified first (exponentially).
Blurring/deblurring of images

original

blurred
We consider the linear problem $Ax \approx b$.

**Existence:** (possible fix) Least squares $A^T Ax = A^T b$ (overdetermined or underdetermined).

**Uniqueness and stability:**

Tikhonov regularization, named after Andrey Tikhonov (1906-1993), is perhaps the most commonly used method of regularization of ill-posed problems.
Tikhonov regularization

One introduces a regularization parameter $\alpha > 0$ in such a way that small $\alpha$ gives us a problem that is "close" to the original. The problem now is to minimize the functional:

$$J_\alpha(x) = \|b - Ax\|^2 + \alpha\|x\|^2,$$

or more generally

$$J_B(x) = \|b - Ax\|^2 + \|Bx\|^2.$$

The solution is given by the linear system $(A^T A + B^T B)x = A^T b$.

$B$ can be a differential operator and $\|Bx\|$ an approximation to a Sobolev norm.
Numerically, the best way to minimize

$$J_B(x) = \| b - Ax \|^2 + \| Bx \|^2.$$

is to solve the concatenated least-squares problem,

$$J_B(x) = \| \begin{bmatrix} A \\ B \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \|^2.$$ 

- For small problems, a QR factorization of $$\begin{bmatrix} A \\ B \end{bmatrix}$$ works well.
- For large (sparse) problems, iterative algorithms can be used.
  - LSQR or LSMR (Paige and Saunders) (click here)
Choice of regularization parameter $\alpha$

This is one of the main topics in the area of inverse problems. The optimal regularization parameter $\alpha$ is usually unknown and often in practical problems is determined by an ad hoc method. Approaches include the discrepancy principle, (leave-one-out) cross-validation, L-curve method, restricted maximum likelihood, unbiased predictive risk estimator, etc.

Morozov’s discrepancy principle (MDP)

A-priori: $\alpha(\delta b) \sim \delta b$

More specifically:

In this principle, $\alpha$ is chosen as the solution of the equation

$$\|Ax_\alpha^\delta - b^\delta\| = C\delta b, \text{ with } C \geq 1.$$
Choice of regularization parameter $\alpha$ (L-curve method)

The L-curve for Tikhonov regularization

Residual norm $\| A x - b \|_2$
Solution norm $\| x \|_2$

$\lambda = 2$
$\| A x - b \|_2 = 6.8574$
$\| x \|_2 = 4.2822$

$\lambda = 0.02$
$\| A x - b \|_2 = 0.062202$
$\| x \|_2 = 7.9112$

$\lambda = 0.003$
$\| A x - b \|_2 = 0.060594$
$\| x \|_2 = 8.2018$

(graphs by P.C. Hansen - click here)
Regularization by SVD filtering

Example

\[ u_t = u_{xx}, \quad -\infty \leq x \leq \infty \quad u(0, x) = f(x), \quad f \in L_2(-\infty, \infty) \]

\[ u(x, t) = \frac{1}{2\sqrt{t\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-\tau)^2}{4t}} f(\tau) d\tau \]

Fredholm first kind integral equation

\[ g(x) = \int_{a}^{b} k(x, \tau)f(\tau) d\tau, \quad a < x, \tau < b \]

When \( k(x, \tau) = k(x - \tau) \). The kernel is spatially invariant.

For image deblurring, typical choice of \( k \)

- Gaussian \( k(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad \sigma > 0 \).

- Out of focus: \( k(x) = \begin{cases} 
C, & x_1 \leq x \leq x_2 \\
0, & \text{otherwise} 
\end{cases} \)
SVD decomposition and filtering

\[ b = Ax_{true} + \eta \]

\[ \delta := \| \eta \| \]

\[ A = U\Sigma V^T \quad \text{invertible} \]

\[ A^{-1} b = V\Sigma^{-1} U^T b = x_{true} + \sum_{i=1}^{n} \sigma_i^{-1} (u_i^T \eta) v_i \]

**Remark:** Instability arises due to division by small singular values.

**Filter:** Multiply \( \sigma_i^{-1} \) by a regularizing filter function \( w_\alpha(\sigma_i^2) \) for which

\[ w_\alpha(\sigma^2)\sigma^{-1} \rightarrow 0 \quad \text{as} \quad \sigma \rightarrow 0. \]

**Regularized Solution**

\[ x_\alpha = \sum_{i=1}^{n} w_\alpha(\sigma_i^2)\sigma_i^{-1} (u_i^T b) v_i. \]
SVD truncation:

\[ w_\alpha(\sigma^2) = \begin{cases} 
1, & \sigma^2 > \alpha \\
0, & \text{otherwise} 
\end{cases} \]

Tikhonov filter: (equivalent to regularization \( \alpha \|x\| \))

\[ w_\alpha(\sigma^2) = \frac{\sigma^2}{\sigma^2 + \alpha} \]

Remark: This can be derived more generally for \textit{compact operators}. 
References

- Prof. Renaut slides:
  http://math.la.asu.edu/ rosie/classes/index.html

  http://www.math.montana.edu/ vogel/Book/

- Rank Deficient and Discrete Ill-Posed Inverse Problems, Hansen, SIAM 1997
  http://www2.imm.dtu.dk/ pch/Regutools/