There and especially back again

An introduction to inverse problems

Yves van Gennip

Technische Universiteit Eindhoven, CASA

21 February 2007

This talk as an inverse problem

Conclusions

- Inverse problems common in science and engineering
- Well-posedness is important
- Concept of worst-case error gives a handle on the error made
- In discrete case, ill-conditioned problems need regularisation
- Order optimality is wanted

Outline

Examples of inverse problems

Polynomials

- Direct problem: find the zeroes of a given polynomial
- Inverse problem: find the polynomial corresponding to given zeroes

Intelligence tests

- DP: evaluate a sequence given its formation law
- IP: construct the formation law of a sequence, given the first few terms

Wave scattering

- DP: given an object and incoming wave, find the scattered wave
- IP: given the far field pattern, find the shape of the object

Examples of inverse problems

Polynomials

- Direct problem: find the zeroes of a given polynomial
- Inverse problem: find the polynomial corresponding to given zeroes

Intelligence tests

- DP: evaluate a sequence given its formation law
- IP: construct the formation law of a sequence, given the first few terms

Wave scattering

- DP: given an object and incoming wave, find the scattered wave
- IP: given the far field pattern, find the shape of the object

Examples of inverse problems

Polynomials

- Direct problem: find the zeroes of a given polynomial
- Inverse problem: find the polynomial corresponding to given zeroes

Intelligence tests

- DP: evaluate a sequence given its formation law
- IP: construct the formation law of a sequence, given the first few terms

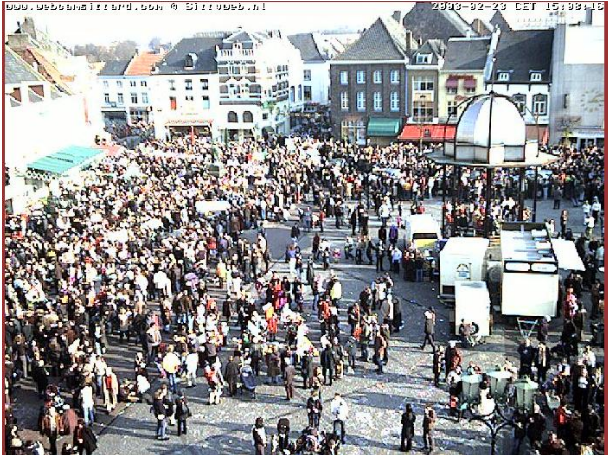
Wave scattering

- DP: given an object and incoming wave, find the scattered wave
- IP: given the far field pattern, find the shape of the object

Sherlock Holmes

Sherlock Holmes looked around and in a part of a second analysed the situation. "Watson - what do you think?" he asked. "Was the murder planned?" - "Well ..." his friend replied. "I have no idea ... how could I know? It seems there is a large mess here, that does not look like a planned event." - "But there are clear signs that all was carefully prepared!" said Sherlock Holmes. This all was a mystery to Watson ... how in the world could he know?

Carnaval



"How was your carnaval?"
"I can't remember, but I have a major hangover."
"Then it must have been good."

General formulation in vector spaces and well-posedness

X and Y are normed spaces, $K : X \to Y$ a mapping

- DP: given $x \in X$ and K, evaluate Kx
- IP: given $y \in Y$ and K, solve $Kx = y$ for $x \in X$

Hadamard's well-posedness

$Kx = y$ is well-posed if:

- Existence: $\forall y \in Y, \exists x \in X, Kx = y$
- Uniqueness: $\forall y (Kx_1 = y \wedge Kx_2 = y \Rightarrow x_1 = x_2)$
- Stability: $\forall (x_n) \subset X$: $Kx_n \to Kx \Rightarrow x_n \to x$

N.B. Stability depends on the topologies of X and Y.

General formulation in vector spaces and well-posedness

X and Y are normed spaces, $K : X \to Y$ a mapping

- DP: given $x \in X$ and K, evaluate Kx
- IP: given $y \in Y$ and K, solve $Kx = y$ for $x \in X$

Hadamard's well-posedness

$Kx = y$ is well-posed if:

- Existence: $\forall y \in Y, \exists x \in X, Kx = y$
- Uniqueness: $\forall y (Kx_1 = y \wedge Kx_2 = y \Rightarrow x_1 = x_2)$
- Stability: $\forall (x_n) \subset X$: $Kx_n \to Kx \Rightarrow x_n \to x$

N.B. Stability depends on the topologies of X and Y.

Example: differentiation

DP: integration

Compute $y(t) = Kx(t) := \int_0^t x(s)\, ds, \quad t \in [0, 1]$

IP: differentiation

- find x solving $Kx = y$, i.e. $x = y'$
- $X = C[0, 1]$ with norm $\|x\|_\infty := \max\limits_{0 \le t \le 1} |x(t)|$
- $Y = \{y \in C[0, 1] : y(0) = 0\}$; choose norm
- $\|\cdot\|_\infty$, ill-posed:

$$\|\delta \sin(t/\delta^2)\|_\infty = \delta, \quad \text{but } \|1/\delta \cos(t/\delta^2)\|_\infty = 1/\delta$$

- $\|y\|_{C^1} := \max\limits_{0 \le t \le 1} |y'(t)|$, well-posed, since K is invertible and bounded:

$$\|Kx\|_{C^1} = \max_{0 \le t \le 1} |x(t)| = \|x\|_\infty$$

Example: differentiation

DP: integration

Compute $y(t) = Kx(t) := \int_0^t x(s)\,ds, \quad t \in [0,1]$

IP: differentiation

- find x solving $Kx = y$, i.e. $x = y'$
- $X = C[0,1]$ with norm $\|x\|_\infty := \max\limits_{0\le t\le 1} |x(t)|$
- $Y = \{y \in C[0,1] : y(0) = 0\}$; choose norm
- $\|\cdot\|_\infty$, ill-posed:

$$\|\delta \sin(t/\delta^2)\|_\infty = \delta, \quad \text{but } \|1/\delta \cos(t/\delta^2)\|_\infty = 1/\delta$$

- $\|y\|_{C^1} := \max\limits_{0\le t\le 1} |y'(t)|$, well-posed, since K is invertible and bounded:

$$\|Kx\|_{C^1} = \max_{0\le t\le 1} |x(t)| = \|x\|_\infty$$

Compact operators

Definition

An operator $K : X \to Y$ is called *compact* if

$$Z \text{ bounded} \Rightarrow \overline{K(Z)} \text{ compact}$$

Example: Fredholm integral equation of the first kind

With continuous kernel k:

$$Kx(t) = \int_a^b k(t-s)x(s)\,ds.$$

Ill-posedness for compact operators

Theorem

- X, Y normed spaces, $K : X \to Y$ linear, compact, nullspace $\mathcal{N}(K)$
- Assume $X/\mathcal{N}(K)$ infinite dimensional
- Then no stability: $\exists (x_n) \subset X$ such that $Kx_n \to 0$, but (x_n) does not converge
- Can choose (x_n) such that $\|x_n\| \to \infty$
- Special case: K one-to-one $\Rightarrow K^{-1}$ unbounded

Sketch of proof

- Construct $\tilde{K} : X/\mathcal{N}(K) \to Y$, compact and one-to-one
- $\tilde{K}^{-1}$ bounded $\Rightarrow I = \tilde{K}^{-1}\tilde{K}$ compact. $\lightning$

Worst-case error

Definition

- X, Y Banach with norms $\|\cdot\|_X$, $\|\cdot\|_Y$
- $K : X \to Y$ linear, bounded
- $X_1 \subset X$ with stronger norm $\|\cdot\|_1$: $\|x\|_X \leq c\|x\|_1$
- $\mathcal{F}(\delta, E, \|\cdot\|_1) := \sup\{\|x\|_X : x \in X_1, \|Kx\|_Y \leq \delta, \|x\|_1 \leq E\}$

Want: $\mathcal{F}(\delta, E, \|\cdot\|_1) = \mathcal{O}(\delta)$

Positive result

K^{-1} bounded: $\|x\|_X \leq \|K^{-1}\|\|Kx\|_Y$

Negative result

- If $K : X \to Y$ linear, compact, $X/\mathcal{N}(K)$ infinite-dimensional
- Then $\forall E > 0, \exists c > 0, \exists \delta_0 > 0 : \forall \delta \in (0, \delta_0), \mathcal{F}(\delta, E, \|\cdot\|_X) \geq c$

Worst-case error

Definition

- X, Y Banach with norms $\|\cdot\|_X$, $\|\cdot\|_Y$
- $K : X \to Y$ linear, bounded
- $X_1 \subset X$ with stronger norm $\|\cdot\|_1$: $\|x\|_X \leq c\|x\|_1$
- $\mathcal{F}(\delta, E, \|\cdot\|_1) := \sup\{\|x\|_X : x \in X_1, \|Kx\|_Y \leq \delta, \|x\|_1 \leq E\}$

Want: $\mathcal{F}(\delta, E, \|\cdot\|_1) = \mathcal{O}(\delta)$

Positive result

K^{-1} bounded: $\|x\|_X \leq \|K^{-1}\|\|Kx\|_Y$

Negative result

- If $K : X \to Y$ linear, compact, $X/\mathcal{N}(K)$ infinite-dimensional
- Then $\forall E > 0, \exists c > 0, \exists \delta_0 > 0 : \forall \delta \in (0, \delta_0), \mathcal{F}(\delta, E, \|\cdot\|_X) \geq c$

Worst-case error

Definition

- X, Y Banach with norms $\|\cdot\|_X$, $\|\cdot\|_Y$
- $K : X \to Y$ linear, bounded
- $X_1 \subset X$ with stronger norm $\|\cdot\|_1$: $\|x\|_X \le c\|x\|_1$
- $\mathcal{F}(\delta, E, \|\cdot\|_1) := \sup\{\|x\|_X : x \in X_1, \|Kx\|_Y \le \delta, \|x\|_1 \le E\}$

Want: $\mathcal{F}(\delta, E, \|\cdot\|_1) = \mathcal{O}(\delta)$

Positive result

K^{-1} bounded: $\|x\|_X \le \|K^{-1}\|\|Kx\|_Y$

Negative result

- If $K : X \to Y$ linear, compact, $X/\mathcal{N}(K)$ infinite-dimensional
- Then $\forall E > 0, \exists c > 0, \exists \delta_0 > 0 : \forall \delta \in (0, \delta_0), \mathcal{F}(\delta, E, \|\cdot\|_X) \ge c$

Example: differentiation

$X = Y = L^2(0,1)$ with $\|\cdot\|_{L^2}$

$Kx(t) = \int_0^t x(s)\,ds, \quad t \in (0,1), x \in L^2(0,1)$

$X_1 := \{x \in H^1(0,1) : x(1) = 0\}$ with $\|x\|_1 := \|x'\|_{L^2}$

- Partial integration and Cauchy-Schwarz: $\|x\|_{L^2}^2 \leq \|Kx\|_{L^2}\|x'\|_{L^2}$
- $\mathcal{F}(\delta, E, \|\cdot\|_1) \leq \sqrt{\delta E}$

$X_2 := \{x \in H^1(0,1) : x(1) = 0, x'(0) = 0\}$ with $\|x\|_2 := \|x''\|_{L^2}$

- Partial integration and C-S: $\|x'\|_{L^2}^2 \leq \|x\|_{L^2}\|x''\|_{L^2}$
- Use X_1-result: $\|x\|_{L^2}^2 \leq \|Kx\|_{L^2}\sqrt{\|x\|_{L^2}}\sqrt{\|x''\|_{L^2}}$
- $\mathcal{F}(\delta, E, \|\cdot\|_2) \leq \delta^{\frac{2}{3}}E^{\frac{1}{3}}$

Example: differentiation

$X = Y = L^2(0,1)$ with $\|\cdot\|_{L^2}$

$Kx(t) = \int_0^t x(s)\,ds, \quad t \in (0,1), x \in L^2(0,1)$

$X_1 := \{x \in H^1(0,1) : x(1) = 0\}$ with $\|x\|_1 := \|x'\|_{L^2}$

- Partial integration and Cauchy-Schwarz: $\|x\|_{L^2}^2 \leq \|Kx\|_{L^2}\|x'\|_{L^2}$
- $\mathcal{F}(\delta, E, \|\cdot\|_1) \leq \sqrt{\delta E}$

$X_2 := \{x \in H^1(0,1) : x(1) = 0, x'(0) = 0\}$ with $\|x\|_2 := \|x''\|_{L^2}$

- Partial integration and C-S: $\|x'\|_{L^2}^2 \leq \|x\|_{L^2}\|x''\|_{L^2}$
- Use X_1-result: $\|x\|_{L^2}^2 \leq \|Kx\|_{L^2}\sqrt{\|x\|_{L^2}}\sqrt{\|x''\|_{L^2}}$
- $\mathcal{F}(\delta, E, \|\cdot\|_2) \leq \delta^{\frac{2}{3}} E^{\frac{1}{3}}$

Example: differentiation

$X = Y = L^2(0,1)$ with $\|\cdot\|_{L^2}$

$Kx(t) = \int_0^t x(s)\,ds, \quad t \in (0,1), x \in L^2(0,1)$

$X_1 := \{x \in H^1(0,1) : x(1) = 0\}$ with $\|x\|_1 := \|x'\|_{L^2}$

- Partial integration and Cauchy-Schwarz: $\|x\|_{L^2}^2 \leq \|Kx\|_{L^2}\|x'\|_{L^2}$
- $\mathcal{F}(\delta, E, \|\cdot\|_1) \leq \sqrt{\delta E}$

$X_2 := \{x \in H^1(0,1) : x(1) = 0, x'(0) = 0\}$ with $\|x\|_2 := \|x''\|_{L^2}$

- Partial integration and C-S: $\|x'\|_{L^2}^2 \leq \|x\|_{L^2}\|x''\|_{L^2}$
- Use X_1-result: $\|x\|_{L^2}^2 \leq \|Kx\|_{L^2}\sqrt{\|x\|_{L^2}}\sqrt{\|x''\|_{L^2}}$
- $\mathcal{F}(\delta, E, \|\cdot\|_2) \leq \delta^{\frac{2}{3}} E^{\frac{1}{3}}$

Two aymptotically sharp estimates

X, Y Hilbert, K linear, compact, one-to-one, $K(X)$ dense

- $X_1 := K^*(Y)$, $\|x\|_1 := \|(K^*)^{-1}x\|_Y \Rightarrow \mathcal{F}(\delta, E, \|\cdot\|_1) \leq \sqrt{\delta E}$
- $X_2 := K^*K(X)$, $\|x\|_2 := \|(K^*K)^{-1}x\|_X \Rightarrow \mathcal{F}(\delta, E, \|\cdot\|_2) \leq \delta^{\frac{2}{3}} E^{\frac{1}{3}}$

Proof uses singular value decomposition (SVD)

SVD, matrix formulation

- X, Y Hilbert, K linear, K^*K has eigenvalues λ_i
- Singular values: $s_i := \sqrt{\lambda_i}$
- $K = U \operatorname{diag}(s_i) V^*$
- Columns of U are orthonormal, columns of V too
- Columns satisfy: $Kv_i = s_i u_i$, $K^*u_i = s_i v_i$

Two aymptotically sharp estimates

X, Y Hilbert, K linear, compact, one-to-one, $K(X)$ dense

- $X_1 := K^*(Y), \|x\|_1 := \|(K^*)^{-1}x\|_Y \Rightarrow \mathcal{F}(\delta, E, \|\cdot\|_1) \leq \sqrt{\delta E}$
- $X_2 := K^*K(X), \|x\|_2 := \|(K^*K)^{-1}x\|_X \Rightarrow \mathcal{F}(\delta, E, \|\cdot\|_2) \leq \delta^{\frac{2}{3}} E^{\frac{1}{3}}$

Proof uses singular value decomposition (SVD)

SVD, matrix formulation

- X, Y Hilbert, K linear, K^*K has eigenvalues λ_i
- Singular values: $s_i := \sqrt{\lambda_i}$
- $K = U \mathrm{diag}(s_i) V^*$
- Columns of U are orthonormal, columns of V too
- Columns satisfy: $Kv_i = s_i u_i, K^* u_i = s_i v_i$

Ill-conditioned discretisations

Discretise and regularise

- Discretised system: $Kf = d$
- Condition number: $\kappa(K) := \|K^{-1}\|\|K\|$, ill-conditioned if κ large
- With L^2-norm: $\kappa(K) = s_{\max}/s_{\min}$
- Error in data: $d = Kf_{\text{true}} + \eta$
- $\delta := \|\eta\| > 0$, $\|\cdot\|$ is Euclidean norm
- Using SVD: $K^{-1}d = f_{\text{true}} + \sum_{i=1}^{n} s_i^{-1}(u_i^T\eta)v_i$
- If s_i is small, s_i^{-1} leads to large error
- Regularising filter function $w_\alpha(s_i^2)$ such that $w_\alpha(s^2)s^{-1} \to 0$ as $s \to 0$
- Approximation of f_{true}: $f_\alpha = \sum_{i=1}^{n} w_\alpha(s_i^2)s_i^{-1}(u_i^T d)v_i$
- Two possible choices considered:
  - Truncate (TSVD): $w_\alpha(s^2) = 1$ if $s^2 > \alpha$, $w_\alpha(s^2) = 0$ otherwise
  - Tikhonov filter: $w_\alpha(s^2) = \frac{s^2}{s^2+\alpha}$

Error analysis and convergence

Error

- Let R_α be such that $f_\alpha = R_\alpha d$
- $e_\alpha := f_\alpha - f_{\text{true}} = e_\alpha^{\text{trunc}} + e_\alpha^{\text{noise}}$
- $e_\alpha^{\text{trunc}} := R_\alpha K f_{\text{true}} - f_{\text{true}}$
- $e_\alpha^{\text{noise}} := R_\alpha \eta$

For which α do errors converge to zero if $\|\eta\| = \delta \to 0$?

- For TSVD and Tikhonov: $e_\alpha^{\text{trunc}} \to 0$ as $\alpha \to 0$
- Both: $\|e_\alpha^{\text{noise}}\| \leq \alpha^{-\frac{1}{2}}\delta$
- Choose $\alpha = \delta^p$ with $0 < p < 2$, then convergence: $e_\alpha \to 0$ as $\delta \to 0$

Order optimality

Order optimality for TSVD

- Assume $\delta \geq s_{\text{min}}^2$
- Assume source condition (or range condition): $f_{\text{true}} = K^T z, z \in \mathbb{R}^n$
- Then sharp inequality: $\|e_\alpha^{\text{trunc}}\|^2 \leq \alpha \|z\|^2$
- Thus: $\|e_\alpha\| \leq \alpha^{\frac{1}{2}} \|z\| + \alpha^{-\frac{1}{2}} \delta$
- Minimise w.r.t. α: $\|e_\alpha\| \leq 2\|z\|^{\frac{1}{2}} \delta^{\frac{1}{2}}$
- Order optimal: $f_{\text{true}} \in \text{Range}(K^T) \Rightarrow \|e_\alpha\| = \mathcal{O}(\sqrt{\delta})$
- N.B. a priori parameter choice rule

A posteriori selection and variational formulation

Moronov's discrepancy principle

- Choose largest α such that $\|Kf_\alpha - d\| \leq \delta$
- TSVD and Tikhonov converge
- Even order optimal

Variational formulation for Tikhonov

- $f_\alpha = \arg\min_{f \in \mathbb{R}^n} \|Kf - d\|^2 + \alpha\|f\|^2$
- Some advantages of this formulation:
  - Incorporate constraints in admissable set, e.g. nonnegativity of f
  - Replace least squares fit $\|Kf - d\|^2$ with other functionals
  - Add extra penalty terms like $\alpha\|f\|^2$, e.g. for a priori information
  - Gradient flow approach (steepest descent) might be possible

A posteriori selection and variational formulation

Moronov's discrepancy principle

- Choose largest α such that $\|Kf_\alpha - d\| \leq \delta$
- TSVD and Tikhonov converge
- Even order optimal

Variational formulation for Tikhonov

- $f_\alpha = \arg\min_{f \in \mathbb{R}^n} \|Kf - d\|^2 + \alpha \|f\|^2$
- Some advantages of this formulation:
  - ▷ Incorporate constraints in admissable set, e.g. nonnegativity of f
  - ▷ Replace least squares fit $\|Kf - d\|^2$ with other functionals
  - ▷ Add extra penalty terms like $\alpha\|f\|^2$, e.g. for a priori information
  - ▷ Gradient flow approach (steepest descent) might be possible

See how well we did

Conclusions

- Examples of inverse problems can be found everywhere
- Often problems are ill-posed: watch the error!
- Given control on the input, the worst-case error can be controlled
- Discussed two regularisation schemes for discrete problems
- TSVD and Tikhonov both are convergent and order optimal

The error looks well within bounds

Starting conclusions

- Inverse problems common in science and engineering
- Well-posedness is important
- Concept of worst-case error gives a handle on the error made
- In discrete case, ill-conditioned problems need regularisation
- Order optimality is wanted

Conclusions after solving inverse talk

- Examples of inverse problems can be found everywhere
- Often problems are ill-posed: watch the error!
- Given control on the input, the worst-case error can be controlled
- Discussed two regularisation schemes for discrete problems
- TSVD and Tikhonov both are convergent and order optimal