When randomness meets inverse problems

Shuai Lu
School of Mathematical Sciences, Fudan University

Seminar RICAM - Fudan, June 10 (Linz-Shanghai), 2020

This talk is based on two recent work

- Shuai Lu, Pingping Niu and Frank Werner: *On the asymptotical regularization for linear inverse problems in presence of white noise,* manuscript

- Shuai Lu, Peter Mathé and Sergei V. Pereverzev: *Randomized matrix approximation to enhance regularized projection schemes in inverse problems,* to appear in Inverse Problems

Outline

Motivation

Continuous regularization for random noise

Steady ill-posed problems to continuous dynamics

Filter Based Methods

Main results

Regularization for random projections

Deterministic inverse problems with random projections

Main results

Conclusion

Motivation

$$y^\delta = Au^\dagger + \eta$$

Where does randomness appear in inverse problems?

- Random noise
- Random operator
- Random solution

Outline

Motivation

Continuous regularization for random noise
- Steady ill-posed problems to continuous dynamics
- Filter Based Methods
- Main results

Regularization for random projections
- Deterministic inverse problems with random projections
- Main results

Conclusion

Linear Inverse Problems

We consider the linear inverse problem

$$y = Au^\dagger + \eta,$$

- A is an injective linear bounded operator acting from a Hilbert space X to Y.
- The exact solution $u^\dagger \in X$.
- Noise $\eta \sim \mathcal{N}(0, \varepsilon^2 I)$.

Regularization schemes

Spectral methods of the form

$$u_\gamma^\delta = q_\gamma(A^*A)A^*y^\delta$$

Variational ones of the form

$$u_\gamma^\delta \in \arg\min_{u\in X}\left[\|Au\|_Y^2 - 2\left\langle Au, y^\delta\right\rangle_{Y\times Y^*} + \gamma R(u)\right]$$

Table 1. Rates for deterministic and stochastic model

	Deterministic	Stochastic
Direct	ε^2	$\varepsilon^{\frac{4\alpha}{2\alpha+1}}$
Polynomial	$\varepsilon^{\frac{4\alpha}{2\alpha+2\beta}}$	$\varepsilon^{\frac{4\alpha}{2\alpha+2\beta+1}}$
Exponential	$\varepsilon^{\frac{4\alpha}{2\alpha+2\beta}}$	$\varepsilon^{\frac{4\alpha}{2\alpha+2\beta}}$

L. Cavalier, *Ch.1 Inverse Problems in Statistics*, P. Alquier et al. (eds.), Inverse Problems and High-Dimensional Estimation, Lecture Notes in Statistics 203, Springer-Verlag Berlin Heidelberg 2011.

Observations

Sequential (and independent) observations

$$y_i = Au^\dagger + \eta_i, \qquad 1 \le i \le N, \quad \eta_i \sim \mathcal{N}_Y(0, \Sigma_i)$$

which yields

$$y^\delta := \frac{1}{N}\sum_{i=1}^{N} y_i = Au^\dagger + \frac{1}{N}\sum_{i=1}^{N}\eta_i, \qquad \frac{1}{N}\sum_{i=1}^{N}\eta_i \sim \mathcal{N}_Y\left(0, \frac{1}{N^2}\sum_{i=1}^{N}\Sigma_i\right).$$

Artificial dynamical system

$$u_n = u_{n-1}$$
$$y_n = Au_n + \eta_n$$

with $\eta_1 \sim \mathcal{N}_Y(0, \Sigma)$.

M. A. Iglesias, K. J. Law, and A. M. Stuart, *Ensemble Kalman methods for inverse problems*, Inverse Problems **29** (2013), no. 4, 045001.

Kalman filter and 3DVAR

The Kalman filter & 3DVAR yield the posterior Gaussian distribution $\mathscr{N}_X(m_n, C_n)$ (or $\mathscr{N}_X(\zeta_n, \mathscr{C})$) where

$$(\text{Kalman filter}) \begin{cases} K_n = C_{n-1}A^*\left(AC_{n-1}A^* + \Sigma\right)^{-1} \\ m_n = m_{n-1} + K_n(y_n - Am_{n-1}) \\ C_n = (I - K_nA)C_{n-1}, \end{cases}$$

$$(\text{3DVAR}) \begin{cases} K_n \equiv \mathscr{K} := C_0A^*\left(AC_0A^* + \Sigma\right)^{-1} \\ \zeta_n = \zeta_{n-1} + \mathscr{K}(y_n - A\zeta_{n-1}), \\ \mathscr{C} \equiv (I - \mathscr{K}A)C_0 \end{cases}$$

with an initial (prior) distribution $\mathscr{N}_X(m_0, C_0)$.

Iglesias, Lin, Lu and Stuart [CMS 2017]; Ding, Lu and Cheng [J. Complexity 2018]

Towards a continuous analog

Artificial dynamical system

$$u_n = u_{n-1}$$
$$y_n = Au_n + \eta_n$$

with $\eta_n \sim \mathcal{N}(0, \tau^{-1}\Sigma)$.

The continuous analog

Denote $z_1, ..., z_n$ be equidistant (approximate) samples of a random process z such that $y_n = \left(\frac{z_n - z_{n-1}}{\tau}\right)$ with a time step $\tau > 0$, let $\tau \to 0$ and

$$\mathrm{d}u = 0, \quad u(0) = u^\dagger;$$
$$\mathrm{d}z = Au\mathrm{d}t + \sqrt{\Sigma}\mathrm{d}W, \quad z(0) = 0.$$

Consistency check

Original steady problem

$$y^{\delta} = Au^{\dagger} + \delta\eta$$

with a noise level δ.

Consistency check

$$y^{\delta} := \frac{1}{T} z(T) = Au^{\dagger} + \frac{1}{T}\sqrt{\Sigma}\left(W(T) - W(0)\right),$$
$$\sqrt{\Sigma}\left(W(T) - W(0)\right) \sim \mathcal{N}_Y(0, T\Sigma).$$

Hence, the ending point of the observable process z carries the same information as the data observed in the original inverse problem with $\boldsymbol{\delta = 1/\sqrt{T}}$.

Kalman-Bucy filter: Non-stationary Asymptotical Regularization Method

The Kalman-Bucy filter:

$$\begin{aligned} \mathrm{d}m &= CA^*\Sigma^{-1}(\mathrm{d}z - Am\mathrm{d}t), \quad m(0) = m_0; \\ \mathrm{d}C &= -CA^*\Sigma^{-1}AC\mathrm{d}t, \quad C(0) = C_0. \end{aligned}$$

Solve the Riccati equation to obtain

$$C(t) = (C_0^{-1} + tA^*\Sigma^{-1}A)^{-1}, \quad t > 0.$$

Thus the Kalman-Bucy filter is

$$\mathrm{d}m = (C_0^{-1} + tA^*\Sigma^{-1}A)^{-1}A^*\Sigma^{-1}(\mathrm{d}z - Am\mathrm{d}t), \quad m(0) = m_0.$$

3DVAR: Stationary Asymptotical Regularization Method

$$\begin{aligned} \mathrm{d}\zeta &= \mathscr{C}A^*\Sigma^{-1}(\mathrm{d}z - A\zeta\mathrm{d}t), \quad \zeta(0) = m_0, \\ \mathrm{d}\mathscr{C} &= 0, \quad \mathscr{C}(0) = C_0. \end{aligned}$$

Non-stationary (and stationary) ARMs

Non-stationary ARM (Kalman-Bucy filter)

$$(u-m)(t)=e^{-\int_0^t (C_0^{-1}+sA^*\Sigma^{-1}A)^{-1}A^*\Sigma^{-1}A\mathrm{d}s}(u-m)(0)$$
$$-\int_0^t e^{-\int_s^t (C_0^{-1}+\tau A^*\Sigma^{-1}A)^{-1}A^*\Sigma^{-1}A\mathrm{d}\tau}(C_0^{-1}+sA^*\Sigma^{-1}A)^{-1}A^*\Sigma^{-1/2}\mathrm{d}W(s).$$

Stationary ARM (3DVAR)

$$(u-\zeta)(t)=e^{-C_0A^*\Sigma^{-1}At}(u-\zeta)(0)-\int_0^t e^{-\int_s^t C_0A^*\Sigma^{-1}A\mathrm{d}\tau}C_0A^*\Sigma^{-1/2}\mathrm{d}W(s).$$

Niu, Lu and Cheng [IPI 2019]

Main ingredients

Effective dimension

$$\mathcal{N}(\gamma) = \mathcal{N}_B(\gamma) := \operatorname{tr}\left((\gamma I + B^*B)^{-1} B^*B\right), \gamma > 0.$$

Stochastic integration (Infinite-dimensional Itô-isometry)

The stochastic integral $\Phi \to \int_0^t \Phi(s) \mathrm{d}\mathscr{W}(s)$ with respect to a Y-valued Q-Wiener process $\mathscr{W}(s)$ satisfies

$$\mathbb{E}\left\| \int_0^t \Phi(s) \mathrm{d}\mathscr{W}(s) \right\|_X^2 = \mathbb{E} \int_0^t \|\Phi(s)\|^2_{\mathscr{L}_2(Y_Q, X)} \mathrm{d}s < \infty$$

for $t \in [0, T]$.

Convergence rate of Non-stationary ARM

Let appropriate Assumptions hold, then the non-stationary ARM yields MSE estimates

- If the function $\lambda \mapsto \phi(\lambda)/\lambda^{p+1}$ is non-increasing, then

$$\mathbb{E}\|m(t)-u^\dagger\|^2 \le \phi^2\left(\left(\frac{\alpha}{t}\right)^{\frac{1}{p+1}}\right) + \alpha^{-\frac{1}{p+1}} t^{-\frac{p}{p+1}} \mathcal{N}\left(\frac{\alpha}{t}\right)$$

  for all $0 \le t \le T$.

- If there is a constant $c < \infty$ with $\phi(\lambda) \le c\lambda^{p+1}$ as $\lambda \to 0$, then

$$\mathbb{E}\|m(t)-u^\dagger\|^2 \le c\left(\frac{\alpha}{t}\right)^2 + \alpha^{-\frac{1}{p+1}} t^{-\frac{p}{p+1}} \mathcal{N}\left(\frac{\alpha}{t}\right)$$

  for all $0 \le t \le T$.

Convergence rate of Stationary ARM

Let appropriate Assumptions hold, then the stationary ARM yields the MSE estimate

$$\mathbb{E}\|\zeta(t)-u^\dagger\|^2 \le c\phi^2\left(\left(\frac{\alpha}{t}\right)^{\frac{1}{p+1}}\right)+\frac{1}{2}\alpha^{-1}\mathrm{tr}\,(\Omega)$$

for all $0\le t\le T$ with a constant $c=\max\left\{(\nu_0/(p+1))^{\nu_0/(p+1)},1\right\}$.

Observation

	Deterministic		Statistical
Steady equations	Tikhonov	$\Longleftrightarrow$	Bayesian approach
Discrete Dynamic	Tikhonov	$\Longleftrightarrow$	Kalman filter
	Iterative Tikhonov	$\Longleftrightarrow$	3DVAR
	Tikhonov	$\Longleftrightarrow$	4DVAR (offline)
Continuous Dynamic	**Tikhonov**	$\Longleftrightarrow$	**Kalman-Bucy filter**
	Showalter	$\Longleftrightarrow$	**3DVAR**

Iglesias, Lin, Lu and Stuart [CMS 2017]; Ding, Lu and Cheng [J. Complexity 2018]; **Lu, Niu, Werner [2020]**

Outline

Linear Inverse Problems

We consider the linear inverse problem

$$y = Ax^{\dagger} + \delta\xi,$$

- A is an injective linear bounded operator acting from a Hilbert space X to Y.
- The exact solution $x^{\dagger} \in X$.
- Noise $\|\xi\|_Y \leq 1$.

Regularized least-squares

Regularized least-squares approximation (RLS):

$$\|Ax - y^\delta\|_Y^2 + \alpha\|x\|_X^2,$$

subject to $x \in X_m = \mathscr{R}(P_m)$.

The parameter $\alpha > 0$ is called the *regularization parameter*. The minimizer is explicitly given as

$$x_\alpha^\delta = (\alpha I + P_m A^* A P_m)^{-1} P_m A^* y^\delta, \quad \alpha > 0.$$

In this case, $n := m$ and Q_m may be regarded as projection onto the range of AP_m. Define $B := AP_m$.

H. W. Engl; M. Hanke and A. Neubauer, *Regularization of inverse problems*. Kluwer Academic Publishers Group, Dordrecht, 1996.

Random projections

Random projections

We let $Q = Q_{Y_{2k}}$ be the orthogonal projection onto the (random) range Y_{2k} of $(BB^*)^q B\Omega$. Given this projection $Q_{Y_{2k}}$ we assign the operator $C = C_{2k,m} := Q_{Y_{2k}} B$ for regularization in the form

$$x_\alpha^\delta := g_\alpha(C^*C)C^*y^\delta = C^* g_\alpha(CC^*)Q_{Y_{2k}} y^\delta,$$

where g_α is a regularization scheme to be introduced later.

Low dimensionality for RLS

For regularized least-squares this is given as $g_\alpha(CC^*) := (\alpha I_{2k} + CC^*)^{-1}$.

N. Halko, P. G. Martinsson, and J. A. Tropp. *Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions*, *SIAM Rev.*, 53(2):217–288, 2011

Convergence analysis

Mean convergence result

Under appropriate assumptions, there is a constant C_R and η_{2k} such that, with index functions $\phi = \theta\psi$

$$\mathbb{E}\|x^\dagger - x_\alpha^\delta\|_X \quad \le C_R\Bigg(\frac{\delta}{\sqrt{\alpha}} + \phi(\alpha) + \theta(\alpha)\left[\psi\left(\mathbb{E}\eta_{2k}^2 + \|(I-Q_n)A\|^2\right)\right] + \mathbb{E}\eta_{2k}^2 + \|(I-Q_n)A\|^2\Bigg).$$

Here we denote the following *approximation theoretic* quantities

$$\eta_{2k} := \|(I - Q_{Y_{2k}})B\|_{X\to Y}.$$

Ingredient for proof:

- Gaussian variant of the Kahane inequality;
- Properties of profile functions;
- The Jensen's inequality...

Outline

When inverse problems meet

- Randomness
- Dynamic systems

Thank you for your attention!

Shuai Lu
School of Mathematical Sciences, Fudan University

Email: slu@fudan.edu.cn
Homepage: http://homepage.fudan.edu.cn/shuailu/